

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:41:19 ; Search time 60 Seconds
(without alignments)
1582.266 Million cell updates/sec

Title: US-10-650-369-12
Perfect score: 1715
Sequence: 1 MVVKVINGEGRIGRLAFRR.....EMSYTAQLVRTLEYFAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1715	100.0	336	5	AAM50665 Streptoco
2	1715	100.0	336	5	AAM50639 Streptoco
3	1714	99.9	336	5	ABP29960 Streptoco
4	1714	99.9	336	6	ABU46455 Protein e
5	1714	99.9	345	5	ABP29106 Streptoco
6	1711	99.8	336	2	AAR56486 Plasmin r
7	1711	99.8	336	4	AAY85681 Streptoco
8	1656.5	96.6	448	5	AAM50664 Streptoco
9	1618	94.3	336	5	AAM50667 Streptoco
10	1618	94.3	336	5	AAM50641 Streptoco
11	1609	93.8	336	5	ABP30758 Streptoco
12	1602	93.4	336	5	AAM50666 Streptoco
13	1602	93.4	336	5	AAM50640 Streptoco
14	1593.5	92.9	335	6	ABU02516 S. pneumo
15	1593.5	92.9	359	4	AAU37576 Streptoco
16	1593.5	92.9	359	6	ABU46262 Protein e
17	1592.5	92.9	359	4	AAU38000 Streptoco
18	1584	92.4	336	5	AAM50669 Streptoco
19	1584	92.4	336	5	AAM50643 Streptoco
20	1579.5	92.1	337	6	ABU44330 Protein e
21	1578	92.0	336	5	AAM50668 Streptoco
22	1578	92.0	336	5	AAM50642 Streptoco
23	1549.5	90.3	333	2	AAW55089 Streptoco
24	1549.5	90.3	333	5	ABP54583 S. pneumo
25	1549.5	90.3	333	7	ADC45135 S. pneumo

26	1410	82.2	336	5	ABB55601	Lactococc
27	1337.5	78.0	333	4	AAU35254	Enterococ
28	1337.5	78.0	333	6	ABU29370	Protein e
29	1333.5	77.8	336	5	ABB53868	Lactococc
30	1325.5	77.3	333	7	ADC95142	E. faeciu
31	1277	74.5	335	6	ABU25342	Protein e
32	1262	73.6	336	5	ABB48810	Listeria
33	1262	73.6	336	6	ABU32568	Protein e
34	1259.5	73.4	334	6	ABU23595	Protein e
35	1217.5	71.0	335	6	ABU43810	Protein e
36	1197.5	69.8	334	6	ABU37737	Protein e
37	1196.5	69.8	357	6	ABP81005	N. gonorr
38	1196.5	69.8	357	6	ABU37077	Protein e
39	1190	69.4	336	4	AAG83058	S. epider
40	1190	69.4	336	6	ABU42966	Protein e
41	1190	69.4	348	5	ABP40668	Staphyloc
42	1183	69.0	336	4	AAU36947	Staphyloc
43	1183	69.0	336	4	AAU34122	Staphyloc
44	1183	69.0	336	6	ABJ19195	Pathogen
45	1183	69.0	336	6	ABU16073	Protein e

ALIGNMENTS

RESULT 1

AAM50665						
ID	AAM50665	standard; protein; 336 AA.				
XX						
AC	AAM50665;					
XX						
DT	08-APR-2002	(first entry)				
XX						
DE	Streptococcus dysgalactiae gapC plasmin binding protein DysGapC.					
XX						
KW	DysGapC; GapC; plasmin binding protein; epitope; infection; vaccine; immunisation; mastitis; therapy.					
KW						
XX	Streptococcus dysgalactiae.					
OS						
XX						
PN	WO200196379-A2.					
XX						
PD	20-DEC-2001.					
XX						
PF	11-JUN-2001; 2001WO-CA0000836.					
XX						
PR	12-JUN-2000; 2000US-0211247P.					
XX						
PA	(UYSA-) UNIV SASKATCHEWAN.					
XX						
PI	Potter AA, Perez-Casal J, Fontaine M;					
XX						
DR	WPI; 2002-098051/13.					
DR	N-PSDB; ABA91328.					
XX						

Novel GapC multiple epitope fusion polypeptide comprising antigenic determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S. parauberis, or S. iniae GapC protein useful for treating mastitis in vertebrates.

Claim 8; Fig 1A-B; 116pp; English.

The present sequence is that of the GapC plasmin binding protein, DysGapC, of Streptococcus dysgalactiae ATCC 43078, an isolate from a case of bovine mastitis. The invention relates to novel GapC multiple epitope fusion proteins that comprise epitopes from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae, Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664). Expression vectors and host cells for production of the multiple epitope fusion proteins are provided. The multiple epitope proteins are used in claimed vaccines for treating or preventing a bacterial infection in a vertebrate, especially a streptococcal infection, and particularly mastitis. They are also used in

CC claimed methods of detecting Streptococcus antibodies. The multiple
CC epitope proteins are capable of eliciting broad immunity against a
CC variety of streptococcal infections while minimising the number of
CC antigens present in the final formulation and concomitantly reducing
CC production costs
XX
SQ Sequence 336 AA;
Query Match 100.0%; Score 1715; DB 5; Length 336;
Best Local Similarity 100.0%; Pred. No. 8.6e-145;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLDPNMLAHLKDYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNSVDEINAAKKAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNSVDEINAAKKAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
RESULT 2
AAM50639
ID AAM50639 standard; protein; 336 AA.
XX
AC AAM50639;
XX
DT 04-APR-2002 (first entry)
XX
DE Streptococcus dysgalactiae gapC plasmin binding protein.
XX
KW GapC; plasmin-binding protein; DysgalGapC; infection; mastitis; vaccine;
KW diagnosis; therapy.
XX
OS Streptococcus dysgalactiae.
XX
PN WO200196381-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000838.
XX
PR 12-JUN-2000; 2000US-0211022P.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
XX
DR WPI; 2002-130725/17.
DR N-PSDB; ABA91248.
XX
PT Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,
PT S.uberis, S.paraberis, or S.iniae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX
PS Claim 1(a); Fig 1A-B; 107pp; English.
XX

CC The present sequence is that of the GapC plasmin binding protein
CC (DysGalGapC) of Streptococcus dysgalactiae ATCC 43078, a clinical isolate
CC from bovine mastitis. It is encoded by the GapC gene given in ABA91248.
CC GapC protein, which has no signal sequence or membrane anchor domain, is
CC capable of eliciting an immune response in a vertebrate. The invention
CC provides the GapC genes and proteins of 5 Streptococcus species, as well
CC as recombinant vectors, host cells and vaccine compositions comprising
CC GapC polynucleotides or proteins. The vaccines are used to treat or
CC prevent a bacterial infection, especially a streptococcal infection, and
CC mastitis in particular (claimed). GapC proteins are also used in claimed
CC methods for detecting GapC antibodies, and to raise antibodies that are
CC used in claimed methods for detecting GapC proteins
XX
SQ Sequence 336 AA;
Query Match 100.0%; Score 1715; DB 5; Length 336;
Best Local Similarity 100.0%; Pred. No. 8.6e-145;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLDPNMLAHLKDYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNSVDEINAAKKAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNSVDEINAAKKAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
RESULT 3
ABP29960
ID ABP29960 standard; protein; 336 AA.
XX
AC ABP29960;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 9096.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX

PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR N-PSDB; ABN70591.
XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
PS Claim 1; Page 4031; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 336 AA;

Query Match 99.9%; Score 1714; DB 5; Length 336;
Best Local Similarity 99.7%; Pred. No. 1.1e-144;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTVFNTNHDILDTGTTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
DB 121 TAPGGNDVKTVFNTNHDILDTGTTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLDKNVSVDENAAKAAASNDSPFGYTEDPIVSSDIVGSYGLFDATQTKVM 300
DB 241 GSVTELVVTLDKNVSVDENAAKAAASNDSPFGYTEDPIVSSDIVGSYGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVTRTLEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVTRTLEYFAKIAK 336

RESULT 4
ABU46455
ID ABU46455 standard; protein; 336 AA.
XX
AC ABU46455;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #31982.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX

OS Streptococcus pyogenes.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US0009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wali D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
DR
DR WPI; 2003-029926/02.
DR N-PSDB; ACA50325.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 74379; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 336 AA;

Query Match 99.9%; Score 1714; DB 6; Length 336;
Best Local Similarity 99.7%; Pred. No. 1.1e-144;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY 121 TAPGNDVKT VVFNTHDILDTGTTTISGASCTTNCCLAPMAKALHDAFGIQKGLMTTIHA 180
Dd 121 TAPGNDVKT VVFNTHDILDTGTTTISGASCTTNCCLAPMAKALHDAFGIQKGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Dd 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Dd 241 GSVTELVVTLTKNSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
Dd 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 5
ABP29106
ID ABP29106 standard; protein; 345 AA.
XX
AC ABP29106;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 7388.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR N-PSDB; ABN69737.
XX

PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
PS Claim 1; Page 3888; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 345 AA;
Query Match 99.9%; Score 1714; DB 5; Length 345;
Best Local Similarity 99.7%; Pred. No. 1.1e-144;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIQLAFRRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Dd 10 MVVKVINGFGRIQLAFRRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 69
QY 61 KEGFVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Dd 70 KEGFVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 129
QY 121 TAPGNDVKT VVFNTHDILDTGTTTISGASCTTNCCLAPMAKALHDAFGIQKGLMTTIHA 180
Dd 130 TAPGNDVKT VVFNTHDILDTGTTTISGASCTTNCCLAPMAKALHDAFGIQKGLMTTIHA 189
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Dd 190 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 249
QY 241 GSVTELVVTLTKNSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Dd 250 GSVTELVVTLTKNSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 309
QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
Dd 310 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 345

RESULT 6
AAR56486
ID AAR56486 standard; protein; 336 AA.
XX
AC AAR56486;
XX
DT 25-MAR-2003 (revised)
DT 22-FEB-1995 (first entry)
XX
DE Plasmin receptor.
XX
KW Plasmin receptor; isolate 64/14; plasmin; alpha-2-antiplasmin;
KW plasminogen activator; bleeding; reocclusion; thrombosis;
KW pulmonary embolism; clots.
XX
OS Streptococcus pyogenes.
XX
PN US5328996-A.
XX
PD 12-JUL-1994.
XX
PF 10-AUG-1992; 92US-00928462.
XX
PR 29-MAR-1989; 89US-00330849.
PR 16-MAY-1990; 90US-00524411.
XX
PA (UYFL) UNIV FLORIDA RES FOUND INC.
XX
PI Von Mering G, Broder C, Boyle MDP, Lottenberg R;
XX WPI; 1994-225327/27.
DR N-PSDB; AAQ70705.
XX
PT New DNA encoding bacterial plasmin receptor - useful as thrombolytic
PT agents, used with plasminogen activator or bound to plasmin, also useful
PT in vaccines against bacterial infection.
XX
PS Claim 1; Col 27-30; 19pp; English.

XX This sequence represents the S. pyogenes plasmin receptor. The DNA
CC encoding this sequence was isolated from the S. pyogenes clinical isolate
CC 64/14. The plasmin receptor has a very high affinity for plasmin which,
CC when bound, retains its enzymatic activity but is not regulated
CC (inhibited) by alpha-2-antiplasmin. The receptor protein, when
CC administered concurrently or sequentially, prolongs the activity of
CC plasminogen activator (2A) so allows a reduction in dose, and thus lowers
CC the risk of bleeding, and may prevent reocclusion of blood vessels. The
CC protein may be coupled to a fibrin-specific monoclonal antibody to
CC provide targeting to clots. The plasmin receptor may be useful in human
CC or veterinary medicine, for treatment of thrombosis and pulmonary
CC embolism, and for solubilising clots in catheters or shunts. (Updated on
CC 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 336 AA;

Query Match 99.8%; Score 1711; DB 2; Length 336;
Best Local Similarity 99.4%; Pred. No. 2e-144;
Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIQLAFRRIONVEGVETVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIQLAFRRIONVEGVETVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVEVNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
DB 121 TAPGNDVKTVEVNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLTKNVSVDENINAAKKAASNDSEGYTEDPIVSSDIVGSYGSLSFDATQTKVM 300
DB 241 GSVTELVTLTKNVSVDENINAAKKAASNDSEGYTEDPIVSSDIVGSYGSLSFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYPAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYPAKIAK 336

RESULT 7
AAAY85681
ID AAY85681 standard; protein; 336 AA.
XX
AC AAY85681;
XX
DT 13-FEB-2001 (first entry)
XX
DE Streptococcal plasmin receptor amino acid sequence.
XX
DE Plasmin receptor; plr; immune response; vaccination; antibacterial;
XX
KW Streptococcal infection; antihelminthic.
XX
OS Streptococcus pyogenes.
XX
PN US6136323-A.
XX
PD 24-OCT-2000.
XX
PF 11-JUL-1994; 94US-00273247.
XX
PR 29-MAR-1989; 89US-00330849.
PR 16-MAY-1990; 90US-00524411.
PR 10-AUG-1992; 92US-00928462.
XX
PA (UYFL) UNIV FLORIDA RES FOUND INC.

PI Von Mering G, Broder C, Lottenberg R, Boyle MDP;
XX
DR WPI; 2001-006210/01.
DR N-PSDB; AAC66140.
XX
PT Raising an immune response in a mammal, especially for vaccination
PT against group A streptococcal infections, comprises administering an
PT isolated streptococcal plasmin receptor protein.
XX
PS Claim 2; Col 27-30; 17pp; English.
XX

CC Polynucleotide sequence AAC66140 encodes the Streptococcal plasmin
CC receptor protein AAY85681. The protein is used in a method for raising an
CC immune response in a mammal. The method comprises administering the
CC plasmin receptor protein. The method is useful as a vaccination against
CC group A Streptococcal infections and potentially against a broad range of
CC infections associated with pathogens expressing glyceraldehyde-3-
CC phosphate dehydrogenase (GAPDH) or GAPDH-like proteins at their cell
CC surfaces. The vaccine has antibacterial and antihelminthic activity
XX

SQ Sequence 336 AA;

Query Match 99.8%; Score 1711; DB 4; Length 336;
Best Local Similarity 99.4%; Pred. No. 2e-144;
Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIQLAFRRIONVEGVETVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIQLAFRRIONVEGVETVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVEVNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
DB 121 TAPGNDVKTVEVNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLTKNVSVDENINAAKKAASNDSEGYTEDPIVSSDIVGSYGSLSFDATQTKVM 300
DB 241 GSVTELVTLTKNVSVDENINAAKKAASNDSEGYTEDPIVSSDIVGSYGSLSFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYPAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYPAKIAK 336

RESULT 8
AAM50664
ID AAM50664 standard; protein; 448 AA.
XX
AC AAM50664;
XX
DT 29-AUG-2003 (revised)
DT 08-APR-2002 (first entry)
XX
DE Streptococcus GapC multiepitope fusion PolyGap4.
XX
KW PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine;
KW immunisation; mastitis; therapy.
XX
OS Streptococcus dysgalactiae.
OS Streptococcus agalactiae.
OS Streptococcus parauberis.
XX
PN WO200196379-A2.
XX
PD 20-DEC-2001.

Db 121 TAPGGDDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180

Qy 181 YTGDMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Db 181 YTGDMILDGPHRGGLRRARAGASINIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Qy 241 GSVTELAVTLTKNSVDENAAKAAASNDSPFVSTEDPIVSSDIVGVSGLFDATQTKVM 300

Db 241 GSVTELAVLEKETSVESINAAKAAASNDSPFVSTEDPIVSSDIIGMAYGSLFDATQTKVQ 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 10

AAM50641

ID AAM50641 standard; protein; 336 AA.

XX AAM50641;

AC AAM50641;

DT 04-APR-2002 (first entry)

XX Streptococcus uberis gapC plasmin binding protein.

DE Streptococcus uberis gapC plasmin binding protein.

DE Streptococcus uberis gapC plasmin binding protein.

KW GapC; plasmin-binding protein; UberGapC; infection; mastitis; vaccine;

KW diagnosis; therapy.

XX Streptococcus uberis.

OS Streptococcus uberis.

XX WO200196381-A2.

PN 20-DEC-2001.

XX 11-JUN-2001; 2001WO-CA000838.

PF 12-JUN-2000; 2000US-0211022P.

PR (UYSA-) UNIV SASKATCHEWAN.

PA Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;

XX WPI; 2002-130725/17.

DR N-PSDB; ABA91250.

XX Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,

PT S.uberis, S.parauberis, or S.iniae, useful as vaccine component for

PT treating streptococcal infection which causes mastitis in vertebrates.

XX Claim 1(c); Fig 3A-B; 107pp; English.

XX The present sequence is that of the GapC plasmin binding protein

CC (UberGapC) of Streptococcus uberis ATCC 9927, encoded by the GapC gene

CC given in ABA91250. GapC, which has no signal sequence or membrane anchor

CC domain, is capable of eliciting an immune response in a vertebrate. The

CC invention provides the GapC genes and proteins of 5 Streptococcus

CC species, as well as recombinant vectors, host cells and vaccine

CC compositions comprising GapC polynucleotides or proteins. The vaccines

CC are used to treat or prevent a bacterial infection, especially a

CC streptococcal infection, and mastitis in particular (claimed). GapC

CC proteins are also used in claimed methods for detecting GapC antibodies,

CC and to raise antibodies that are used in claimed methods for detecting

CC GapC proteins. S. uberis is a common pathogen associated with mastitis in

CC cattle, horse, sheep and goat

XX Sequence 336 AA;

SQ Query Match

Best Local Similarity 94.3%; Score 1618; DB 5; Length 336;

Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MVVKVINGFGRIGRLAFRRIQNVGVEVTRINDLTDNMLAHLKDYDTTQGRFDGTVEV 60

Db 1 MVVKVINGFGRIGRLAFRRIQNVGVEVTRINDLTDNMLAHLKDYDTTQGRFDGTVEV 60

Qy 61 KEGGFEVNGNFIKVSABERDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120

Db 61 KEGGFEVNGNFIKVSABERDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120

Qy 121 TAPGGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180

Db 121 TAPGGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180

Qy 181 YTGDMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Db 181 YTGDMILDGPHRGGLRRARAGASINIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Qy 241 GSVTELAVTLTKNSVDENAAKAAASNDSPFVSTEDPIVSSDIVGVSGLFDATQTKVM 300

Db 241 GSVTELAVLEKETSVESINAAKAAASNDSPFVSTEDPIVSSDIIGMAYGSLFDATQTKVQ 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 11

ABP30758

ID ABP30758 standard; protein; 336 AA.

XX ABP30758;

AC ABP30758;

XX 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 10692.

DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;

KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.

OS WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

PF 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;

PI Tettelin H;

XX WPI; 2002-352536/38.

DR N-PSDB; ABN71389.

XX New Streptococcus protein for the treatment or prevention of infection or

PT disease caused by Streptococcus bacteria, such as meningitis, and for

PT detecting a compound that binds to the protein.

XX Claim 1; Page 4179; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

CC the specification. The proteins have antibacterial and antiinflammatory

CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 336 AA;

Query Match 93.8%; Score 1609; DB 5; Length 336;
Best Local Similarity 92.0%; Pred. No. 2.6e-135;
Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 60

Qy 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGFEVNGQFVKVSAEREPANIDWATDGVIEIVLEATGFFASKEAEQHENGAKKVI 120

Qy 121 TAPGNDVKTVVFNTHDILDTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 121 TAPGNDVKTVVFNTHDILDTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180

Qy 181 YTGDMILDGPHRGDLRRARAGAAANIPVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAANIPVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Qy 241 GSVTELVTLDKNSVDENINAAKASNDISFGYTEDPIVSSDIVGSYGLFDATQTKVM 300
Db 241 GSVTELVTLEKDVTEEVAAMKAAANDSYGYTEDPIVSSDIVGSYGLFDATQTKVQ 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSLVRLTLEYFAKIAK 336

RESULT 12
AAM50666
ID AAM50666 standard; protein; 336 AA.
XX
AC AAM50666;
XX
DT 08-APR-2002 (first entry)
XX
DE Streptococcus agalactiae gapC plasmin binding protein AgalGapC.
XX
KW AgalGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
KW immunisation; mastitis; therapy.
XX
OS Streptococcus agalactiae.
XX
PN WO200196379-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000836.
XX
PR 12-JUN-2000; 2000US-0211247P.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Potter AA, Perez-Casal J, Fontaine M;
XX
DR WPI; 2002-098051/13.
DR N-PSDB; ABA91329.
XX
PT Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in

vertebrates.
XX
PS Claim 8; Fig 2A-B; 116pp; English.
XX
CC The present sequence is that of the GapC plasmin binding protein,
CC AgalGapC, of Streptococcus agalactiae ATCC 27541. The invention relates
CC to novel GapC multiple epitope fusion proteins that comprise epitopes
CC from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
CC (see AAM50665-69). A claimed example is polyGap4 (see AAM50664).
CC Expression vectors and host cells for production of the multiple epitope
CC fusion proteins are provided. The multiple epitope proteins are used in
CC claimed vaccines for treating or preventing a bacterial infection in a
CC vertebrate, especially a streptococcal infection, and particularly
CC mastitis. They are also used in claimed methods of detecting
CC Streptococcus antibodies. The multiple epitope proteins are capable of
CC eliciting broad immunity against a variety of streptococcal infections
CC while minimising the number of antigens present in the final formulation
CC and concomitantly reducing production costs
XX
SQ Sequence 336 AA;

Query Match 93.4%; Score 1602; DB 5; Length 336;
Best Local Similarity 91.7%; Pred. No. 1.1e-134;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 60

Qy 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGFEVNGQFVKVSAEREPANIDWATDGVIEIVLEATGFFASKEAEQHENGAKKVI 120

Qy 121 TAPGNDVKTVVFNTHDILDTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 121 TAPGNDVKTVVFNTHDILDTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180

Qy 181 YTGDMILDGPHRGDLRRARAGAAANIPVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAANIPVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Qy 241 GSVTELVTLDKNSVDENINAAKASNDISFGYTEDPIVSSDIVGSYGLFDATQTKVM 300
Db 241 GSVTELVTLEKDVTEEVAAMKAAANDSYGYTEDPIVSSDIVGSYGLFDATQTKVQ 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSLVRLTLEYFAKIAK 336

RESULT 13
AAM50640
ID AAM50640 standard; protein; 336 AA.
XX
AC AAM50640;
XX
DT 04-APR-2002 (first entry)
XX
DE Streptococcus agalactiae gapC plasmin binding protein.
XX
KW GapC; plasmin-binding protein; AgalGapC; infection; mastitis; vaccine;
KW diagnosis; therapy.
XX
OS Streptococcus agalactiae.
XX
PN WO200196381-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000838.
XX
PR 12-JUN-2000; 2000US-0211022P.

XX (UYSA-) UNIV SASKATCHEWAN.
PA Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
XX WPI; 2002-130725/17.
DR N-PSDB; ABA91249.
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,
PT S.uberis, S.parauheris, or S.iniae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX Claim 1(b); Fig 2A-B; 107pp; English.
PS The present sequence is that of the GapC plasmin binding protein
XX (AgalGapC) of Streptococcus agalactiae ATCC 27541, encoded by the GapC
CC gene given in ABA91249. GapC, which has no signal sequence or membrane
CC anchor domain, is capable of eliciting an immune response in a
CC vertebrate. The invention provides the GapC genes and proteins of 5
CC Streptococcus species, as well as recombinant vectors, host cells and
CC vaccine compositions comprising GapC polynucleotides or proteins. The
CC vaccines are used to treat or prevent a bacterial infection, especially a
CC streptococcal infection, and mastitis in particular (claimed). GapC
CC proteins are also used in claimed methods for detecting GapC antibodies,
CC and to raise antibodies that are used in claimed methods for detecting
CC GapC proteins. S. agalactiae is a common pathogen associated with
CC mastitis in cattle, horse, sheep and goat. It also causes septicemia,
CC meningitis, bacteraemia, impetigo, arthritis, urinary tract infections,
CC abscesses, spontaneous abortion, etc
XX Sequence 336 AA;
SQ

Query Match 93.4%; Score 1602; DB 5; Length 336;
Best Local Similarity 91.7%; Pred. No. 1.1e-134;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRICRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRICRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 120
Db 61 KEGGFEVNGQFVKVSAEREPANIDWATDGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 120
QY 121 TAPGNDVKTVTNTNHDILDTGTETVIGASCTTNCLAPMAKALHDAFGIOKGLMTTIIA 180
Db 121 TAPGNDVKTVTNTNHDILDTGTETVIGASCTTNCLAPMAKALQDNFVGVKQGLMTTIIA 180
QY 181 YTGDMILGPHRGDLRRARAGANIVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILGPHRGDLRRARAGANIVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLDKNSVDEINAAKKAANDSPGYTDPVSSDIVGSYSLFDTATQTKVM 300
Db 241 GSVTELVTLEKDVTEEVNAAKKAANDSPGYTDPVSSDIVGSYSLFDTATQTKVQ 300
QY 301 EVDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSVYDNEMSYTSQLVRLTLEYFAKIAK 336

RESULT 14
ABU02516
ID ABU02516 standard; protein; 335 AA.
XX ABU02516;
AC
XX 23-OCT-2003 (revised)
DT 11-FEB-2003 (first entry)
XX S. pneumoniae type 4 strain protein from coding region #2094.
DE
XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;

KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
XX gene therapy; vaccine.
OS Streptococcus pneumoniae; type 4 strain.
XX WO200277021-A2.
PN 03-OCT-2002.
XX 27-MAR-2002; 2002WO-IB002163.
PF 27-MAR-2001; 2001GB-00007658.
PR (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
PA Masignani V, Tettelin H, Fraser C;
XX WPI; 2003-040579/03.
DR N-PSDB; ABX07806.
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT ear infection.
XX Claim 1; SEQ ID NO 4188; 56pp; English.
XX The invention relates to a protein comprising or having at least 50%
CC identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC ABS56454. Also included are an antibody which binds one of the proteins,
CC treating a patient by administering the protein, DNA or antibody (in a
CC composition), a kit comprising first and second primers, which are the
CC nucleic acid cited above or fragments between nucleotides 8-100 of a
CC sequence not defined in the specification, for amplifying a target
CC sequence contained within a Streptococcus nucleic acid sequence, where
CC the first primer is substantially complementary to the target sequence
CC and the second primer is substantially complementary to the complement of
CC the target sequence, and where the parts of the primers having
CC substantial complementarity define the termini of the target sequence to
CC be amplified, assay comprising contacting a test compound with the
CC protein, and determining whether the test compound binds to the protein
CC and a Streptococcus pneumoniae bacterium, where one or more genes
CC encoding the proteins has been rendered inactive. The proteins, nucleic
CC acid molecules, antibody and compositions are useful as medicaments for
CC treating or preventing a disease or infection due to streptococcus
CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC media or ear infection. They are also useful in developing vaccines,
CC diagnostics and antibiotics. The methods are useful for identifying
CC immunodominant proteins. The present sequence is one of the 2469 proteins
CC expressed by the identified coding regions from the genomic sequence.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
CC standardise OS field)

XX Sequence 335 AA;

Query Match 92.9%; Score 1593.5; DB 6; Length 335;
Best Local Similarity 92.6%; Pred. No. 6.3e-134;
Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;
QY 1 MVVKVINGFGRICRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRICRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 120
Db 61 KEGGFEVNGKFIKVSARDPENIDWATDGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 119

QY	121	TAPGGNDVKTVVFNTHDILDGTTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIIHA	180
Dd	120	TAPGGNDVKTVVFNTHDVLDTGTETVISGASCTTNCLAPMAKALQDNFVGVEGLMTTIIHA	179
QY	181	YTGDQMILDGPHRGGDLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGAQRVPVPT	240
Dd	180	YTGDQMILDGPHRGGDLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGSAQRVPTPT	239
QY	241	GSVTTELVTLLDKNVSVDENAAKKAASNDSEFYTEDPVISSDIVGVSYGSLFDATQTQVM	300
Dd	240	GSVTTELVAVLKNVTIDEVNAAMKAASNESGYTEDPVISSDIVGMSYGSLFDATQTQVL	299
QY	301	EVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIAK	336
Dd	300	DVDGKOLVKVVSWYDNEMSYTAQLVRTLEYFAKIAK	335

RESULT 15

RESUL ID
AAU37576
ID AAU37576 standard; protein: 359 AA.

ID
XX
AC

XX
DT 14-FEB-2002 (first entry)

XX Streptococcus pneumoniae cellular proliferation protein #5.

XX
streptococcus pneumoniae

AA Antisense: prok

KW Anticubense, p
KW antibacteria

[illegible]

OS Streptococcus pneumoniae.

XXXXXX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

21-MAR-2001; 2001WO-US0009180.

XX
21 10 2000 - 2000 07 01

PR	21-MAR-2008; 2000US=01910
PR	23-MAY-2008; 2000US=02068

PR 23-MAY-2000; 2000
PR 26-MAY-2000: 2000

PR 26-MAY-2000; 2000US-020772/E;
PR 23-OCT-2000; 2000US-0242578P;

PR 27-NOV-2000: 2000US-0253625P.

22-DEC-2000; 2000US-0257931P.
16-FEB-2001; 2001US-0269308P.
(ELIT-) ELITRA PHARM INC.
Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Yamamoto RT, Xu HH;
WPI; 2001-611495/70.
N-PSDB; AAS55435.
New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids.
Example 3; SEQ ID NO 13169; 511pp; English.
The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are *Escherichia*
coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
Pseudomonas aeruginosa and *Enterococcus faecalis*. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence represents an essential prokaryotic

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:46:47 ; Search time 23 Seconds
(without alignments)
754.188 Million cell updates/sec

Title: US-10-650-369-12
Perfect score: 1715
Sequence: 1 MVVKVINGFGRIGRLAFRR.....EMSYTAQLVRLLEYFAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1715	100.0	336	4	US-09-878-766A-12
2	1711	99.8	336	1	US-07-928-462-2
3	1711	99.8	336	3	US-08-273-247-2
4	1656.5	96.6	448	4	US-09-878-766A-22
5	1618	94.3	336	4	US-09-878-766A-16
6	1602	93.4	336	4	US-09-878-766A-14
7	1584	92.4	336	4	US-09-878-766A-20
8	1578	92.0	336	4	US-09-878-766A-18
9	1549.5	90.3	333	3	US-08-961-083-54
10	1549.5	90.3	333	4	US-09-535-784-54
11	1337.5	78.0	346	4	US-09-134-000C-4400
12	1325.5	77.3	333	4	US-09-107-532A-4769
13	1190	69.4	348	4	US-09-134-001C-5513
14	967	56.4	357	4	US-09-134-000C-4229
15	966	56.3	340	4	US-09-634-238-233
16	959	55.9	349	4	US-09-489-039A-13939
17	896.5	52.3	335	6	5290690-9
18	885.5	51.6	334	6	5290690-10
19	835.5	48.7	334	6	5290690-11
20	830	48.4	335	4	US-09-489-039A-7679
21	824.5	48.1	340	3	US-09-095-855-205
22	824.5	48.1	340	4	US-09-205-426-205
23	799	46.6	340	4	US-09-543-681A-8083
24	770.5	44.9	352	4	US-09-489-039A-9410
25	766.5	44.7	282	4	US-09-107-532A-4057
26	740.5	43.2	336	4	US-09-198-452A-662
27	736	42.9	338	4	US-09-091-725-10

28	718.5	41.9	335	2	US-08-903-800A-6	Sequence 6, Appli
29	715	41.7	340	4	US-09-252-991A-21612	Sequence 21612, A
30	709.5	41.4	334	4	US-09-674-826B-6	Sequence 6, Appli
31	699.5	40.8	268	2	US-08-997-080-186	Sequence 186, App
32	699.5	40.8	268	2	US-08-997-362-186	Sequence 186, App
33	699.5	40.8	268	3	US-09-095-855-186	Sequence 186, App
34	699.5	40.8	268	4	US-09-324-542-186	Sequence 186, App
35	699.5	40.8	268	4	US-09-205-426-186	Sequence 186, App
36	697.5	40.7	341	4	US-09-134-001C-3783	Sequence 3783, Ap
37	667.5	38.9	333	3	US-09-532-803-11	Sequence 11, Appli
38	659.5	38.5	333	3	US-09-532-803-9	Sequence 9, Appli
39	645	37.6	277	4	US-09-548-938A-13	Sequence 13, Appli
40	580	33.8	350	4	US-09-489-039A-11854	Sequence 11854, A
41	572.5	33.4	340	4	US-09-328-352-6704	Sequence 6704, Ap
42	570	33.2	442	4	US-09-540-236-3378	Sequence 3378, Ap
43	553	32.2	417	4	US-09-540-236-3470	Sequence 3470, Ap
44	552	32.2	359	4	US-09-543-681A-6082	Sequence 6082, Ap
45	549	32.0	367	4	US-09-252-991A-19846	Sequence 19846, A

ALIGNMENTS

RESULT 1
US-09-878-766A-12
; Sequence 12, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-766A-12

Query Match	100.0%	Score 1715;	DB 4;	Length 336;
Best Local Similarity	100.0%	Pred. No. 4.2e-164;		
Matches	336;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV	60	
Db	1	MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV	60	
QY	61	KEGGFEVNGNFIKVSARDPENIDWATDGVETIVLEATGFFAKKEAEKHLHANGAKKVI	120	
Db	61	KEGGFEVNGNFIKVSARDPENIDWATDGVETIVLEATGFFAKKEAEKHLHANGAKKVI	120	
QY	121	TAPGGNDVKTIVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA	180	
Db	121	TAPGGNDVKTIVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA	180	
QY	181	YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT	240	
Db	181	YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT	240	
QY	241	GSVTELVVTLDKNVSVDEINAAKKAANDSFGYTEDPIVSSDIVGVYGS�FDATQTQKVM	300	
Db	241	GSVTELVVTLDKNVSVDEINAAKKAANDSFGYTEDPIVSSDIVGVYGS�FDATQTQKVM	300	
QY	301	EVDGSQLVKVSVWYDNEMSYTAQLVRLLEYFAKIAK	336	
Db	301	EVDGSQLVKVSVWYDNEMSYTAQLVRLLEYFAKIAK	336	

RESULT 2
US-07-928-462-2
; Sequence 2, Application US/07928462
; Patent No. 5328996
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmin Receptors as
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,462
; FILING DATE: 19920810
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-462-2

Query Match 99.8%; Score 1711; DB 1; Length 336;
Best Local Similarity 99.4%; Pred. No. 1.1e-163;
Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVVNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 121 TAPGNDVKTVVNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLDKNVSVDEINAAKKAASNDSFGYTEDPIVSSDIVGSYSLFDATQTKVM 300
Db 241 GSVTELVTLDKNVSVDEINAAKKAASNDSFGYTEDPIVSSDIVGSYSLFDATQTKVM 300

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
RESULT 3
US-08-273-247-2
; Sequence 2, Application US/08273247
; Patent No. 6136323
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmin Receptors as
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,247
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-273-247-2

Query Match 99.8%; Score 1711; DB 3; Length 336;
Best Local Similarity 99.4%; Pred. No. 1.1e-163;
Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLTDPNMLAHLKLYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVVNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 121 TAPGNDVKTVVNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

QY 241 GSVTELVVTLTKNVSVDENAAKKAASNDSEFGYTEDPIVSSDIVGVSYSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDENAAKKAASNDSEFGYTEDPIVSSDIVGVSYSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 4

US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

Query Match 96.6%; Score 1656.5; DB 4; Length 448;
Best Local Similarity 79.6%; Pred. No. 5.1e-158;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 28 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
QY 61 KEGGFVNGNFIKVSARDPE-----NIDWATDGVLEIVLE 81
Db 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEIVLEALEGTVEVKDGGFDVNGKFIKVSAR 147
QY 82 -----NIDWATDGVLEIVLE 95
Db 148 KDEQIDWATDGVLEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLE 207
QY 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDGTETVSGASCTTN 155
Db 208 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDGTETVSGASCTTN 267
QY 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA 215
Db 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA 327
QY 216 KAIGLVIPELNGKLDGAAQRPVPVPTGSVTELVVTLTKNVSVDENAAKKAASNDSEFGYTE 275
Db 328 KAIGLVIPELNGKLDGAAQRPVPVPTGSVTELVVTLTKNVSVDENAAKKAASNDSEFGYTE 387
QY 276 DPIVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
Db 388 DPIVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
QY 336 K 336
Db 448 K 448

RESULT 5

US-09-878-766A-16

; Sequence 16, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match 94.3%; Score 1618; DB 4; Length 336;
Best Local Similarity 92.9%; Pred. No. 2.4e-154;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVLEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSARDPENIDWATDGVLEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTVPNTNHDILDGTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHA 180
Db 121 TAPGGNDVKTVPNTNHDILDGTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPPT 240
QY 241 GSVTELVVTLTKNVSVDENAAKKAASNDSEFGYTEDPIVSSDIVGVSYSGLFDATQTKVM 300
Db 241 GSVTELVVLEKETSVEEINAAKKAASNDSEFGYTEDPIVSSDIIGMAYGSLFDATQTKVQ 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 6

US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match 93.4%; Score 1602; DB 4; Length 336;
Best Local Similarity 91.7%; Pred. No. 9.9e-153;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGEVTRINDLTDNMLAHLKLYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGEVTRINDLTDNMLAHLKLYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSABERDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEATGFFASKKAGQHIHENGAKKVI 120
QY 121 TAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
DB 121 TAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLDKNVSVDEINAAKAAASNDSEGYTEDPIVSSDIVGSLSLFDATQTKVM 300
DB 241 GSVTELVTLEKDVTVVEVNAAMKAAANDSYGYTEDPIVSSDIVGSLSLFDATQTKVQ 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
DB 301 TVDGNQLVKVSWYDNEMSYTSQLVRTLEYFAKIAK 336

RESULT 7

US-09-878-766A-20
; Sequence 20, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus iniae
US-09-878-766A-20

Query Match 92.4%; Score 1584; DB 4; Length 336;
Best Local Similarity 90.5%; Pred. No. 6.4e-151;
Matches 304; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGEVTRINDLTDNMLAHLKLYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGEVTRINDLTDNMLAHLKLYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSABERDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGSFVKVSAEREPANIDWATDGVVDIVLEATGFFASKAAAEQHIHANGAKKVI 120
QY 121 TAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
DB 121 TAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHG 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMVLDGPHRGDLRRARAGAAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLDKNVSVDEINAAKAAASNDSEGYTEDPIVSSDIVGSLSLFDATQTKVM 300
DB 241 GSVTELVALEKDVTVVEVNAAMKAAANDSYGYTEDPIVSSDIVGSLSLFDATQTKVQ 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
DB 301 TVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 8

US-09-878-766A-18
; Sequence 18, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus parauberis
US-09-878-766A-18

Query Match 92.0%; Score 1578; DB 4; Length 336;
Best Local Similarity 90.8%; Pred. No. 2.6e-150;
Matches 305; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGEVTRINDLTDNMLAHLKLYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGEVTRINDLTDNMLAHLKLYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSABERDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KGGFDVNGKFIKVSABERDPENIDWATDGVVEIVLEATGFFAKKAAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
DB 121 TAPGGDDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLDKNVSVDEINAAKAAASNDSEGYTEDPIVSSDIVGSLSLFDATQTKVM 300
DB 241 GSVTELVAVLNKETSVEEINSMKAAANDSYGYTEDPIVSSDIVGMSFSLFDATQTKVQ 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
DB 301 TVDGNQLVKVSWYDNEMSYTAQLDRRTLEYFAKIAK 336

RESULT 9

US-08-961-083-54
; Sequence 54, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/961,083
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brookes, A. Anders
/ REGISTRATION NUMBER: 36,373
/ REFERENCE/DOCKET NUMBER: PB340P2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 54:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 333 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-961-083-54

Query Match 90.3%; Score 1549.5; DB 3; Length 333;
Best Local Similarity 91.3%; Pred. No. 1.9e-147;
Matches 303; Conservative 12; Mismatches 16; Indels 1; Gaps 1;

```
QY 2 VVKVGINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEVK 61
Db 1 VVKVGINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEVK 60

QY 62 EGGFEVNGNFIKVSABERDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVVIT 121
Db 61 EGGFEVNGKFKIVSABERDPEQIDWATDGVETVLEATGFFAKKEAAEKHL-KGGAKKVVIT 119

QY 122 APGGNDVKTIVFNTNHDILDTGTTETVISGASCTTNCCLAPMAKALHDAFGIOKGLMTTIHAY 181
Db 120 APGGNDVKTIVFNTNHDVLDGTETVISGASCTTNCCLAPMAKALQDNFVVEGLMTTIHAY 179

QY 182 TGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTG 241
Db 180 TGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGSAQRPVPTG 239

QY 242 SVTELVTLDKNSVDEINAAKASNDSPGYTEDPIVSSDIVGVSGLSFDATQTKVME 301
Db 240 SVTELVAVLEKNVTVDENVAAKASNESYGYTEDPIVSSDIVGMSYGLSFDATQTKVLD 299

QY 302 VDGSQLVKVSWSYDNEMSYTAQLVRTLEYFAK 333
Db 300 VDGSQLVKVSWSYDNEMSYTAQLVRTLGILRK 331
```

RESULT 10
US-09-536-784-54
; Sequence 54, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784

/ FILING DATE: 30-Oct-1997
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/961,083
/ FILING DATE: OCT-30-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Michelle S. Marks
/ REGISTRATION NUMBER: 41,971
/ REFERENCE/DOCKET NUMBER: PB340P3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 54:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 333 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 54:
/ US-09-536-784-54

Query Match 90.3%; Score 1549.5; DB 4; Length 333;
Best Local Similarity 91.3%; Pred. No. 1.9e-147;
Matches 303; Conservative 12; Mismatches 16; Indels 1; Gaps 1;

```
QY 2 VVKVGINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEVK 61
Db 1 VVKVGINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEVK 60

QY 62 EGGFEVNGNFIKVSABERDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVVIT 121
Db 61 EGGFEVNGKFKIVSABERDPEQIDWATDGVETVLEATGFFAKKEAAEKHL-KGGAKKVVIT 119

QY 122 APGGNDVKTIVFNTNHDILDTGTTETVISGASCTTNCCLAPMAKALHDAFGIOKGLMTTIHAY 181
Db 120 APGGNDVKTIVFNTNHDVLDGTETVISGASCTTNCCLAPMAKALQDNFVVEGLMTTIHAY 179

QY 182 TGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTG 241
Db 180 TGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGSAQRPVPTG 239

QY 242 SVTELVTLDKNSVDEINAAKASNDSPGYTEDPIVSSDIVGVSGLSFDATQTKVME 301
Db 240 SVTELVAVLEKNVTVDENVAAKASNESYGYTEDPIVSSDIVGMSYGLSFDATQTKVLD 299

QY 302 VDGSQLVKVSWSYDNEMSYTAQLVRTLEYFAK 333
Db 300 VDGSQLVKVSWSYDNEMSYTAQLVRTLGILRK 331
```

RESULT 11
US-09-134-000C-4400
; Sequence 4400, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4400
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-4400
Query Match 78.0%; Score 1337.5; DB 4; Length 346;

Best Local Similarity 78.4%; Pred. No. 4.1e-126;
Matches 262; Conservative 24; Mismatches 47; Indels 1; Gaps 1;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRNLDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 14 MTVKVINGFGRIGRLAFRRIONVEGVETRNLDLTPNMLAHLKLYDTTQGRFDGTVEV 73
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 74 HEGSFVNGKEIKVLANRNPPELPGELGVDIVLECTGFFTSKSAAEKHLTA-GAKRVVI 132
QY 121 TAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 133 SAPGGNDVPTIVYNTNHTLTGEETVISGASCTTNCLAPMAKALHDFGVVEGLMTTIHA 192
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 193 YTGDMILDGPHPKGDFRRARAAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVAT 252
QY 241 GSVTELVTLDKXNSVDEINAAKAAASNDSPGYTEDPIVSSDIVGSYSLFDATQTKVM 300
Db 253 GSLTELVTLDKEVTVDVNAVMKAAANESYGYNTDEIVSSDIVGMYGSLFDATQTKVM 312
QY 301 EVDGSQLVKSVDNEMSYTAQLVRLTLEYFAKI 334
Db 313 TVGDKQLVKTVAWYDNEMSYTAQLVRLTLEYFANL 346
RESULT 12
US-09-107-532A-4769
; Sequence 4769, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4769:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium

FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...333
; SEQUENCE DESCRIPTION: SEQ ID NO: 4769:
US-09-107-532A-4769
Query Match 77.3%; Score 1325.5; DB 4; Length 333;
Best Local Similarity 77.8%; Pred. No. 6.3e-125;
Matches 260; Conservative 25; Mismatches 48; Indels 1; Gaps 1;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRNLDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MTVKVINGFGRIGRLAFRRIONVEGVETRNLDLTPNMLAHLKLYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 HEGSFVNGKEIKVLANRNPPELPGELGVDIVLECTGFFTSKSAAEKHLTA-GAKRVVI 119
QY 121 TAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 120 SAPGGNDVPTIVYNTNHTLTGEETVISGASCTTNCLAPMAKALHDFGVVEGLMTTIHA 179
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 180 YTGDMILDGPHPKGDFRRARAAAANIVPNSTGAAGAIGLVIPELNGKLYGAAQRPVPT 239
QY 241 GSVTELVTLDKXNSVDEINAAKAAASNDSPGYTEDPIVSSDIVGSYSLFDATQTKVM 300
Db 240 GSLTELVTLEKEVTVDVNAVMKAAANESYGYNTDEIVSSDIVGMYGSLFDATQTKVM 299
QY 301 EVDGSQLVKSVDNEMSYTAQLVRLTLEYFAKI 334
Db 300 TVGDKQLVKTVAWYDNEMSYTAQLVRLTLEYFANL 333
RESULT 13
US-09-134-001C-5513
; Sequence 5513, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5513
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5513
Query Match 69.4%; Score 1190; DB 4; Length 348;
Best Local Similarity 69.4%; Pred. No. 2.9e-111;
Matches 234; Conservative 39; Mismatches 62; Indels 2; Gaps 2;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRNLDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 13 MAIKVAINGFGRIGRLAFRRIONVEGVETRNLDLTPNMLAHLKLYDTTQGRFDGTVEV 72
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 73 IEGGFVNGKEIKSFEDEPDAGKLPWGDLDDIVVLECTGFTDKEKAQAHAIDA-GAKKVI 131
QY 121 TAPGGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 132 SAPAKGDVKTIVFNTHDLDGSETVVSASCTTNSLAPVAKVLSDEFGLVEGFMPTIHA 191


```
QY 181 YTGDMILDPHRRGDLRRARAGAAIVNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 192 YTGDMILDPHRRGDLRRARAGAAIVNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 251
QY 241 GSVTELVTLK-NVSVDEINAAKASNDSTGYTEDPIVSSDIVGVSGLFDATQKV 299
Db 252 GSVTELVTLK-NVSVDEINAAKASNDSTGYTEDPIVSSDIVGVSGLFDATQKV 311
QY 300 MEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 312 MTVGDRQLVKVAAWYDNEMSYTAQLVRLTLEYFAKIAK 348

RESULT 14
US-09-134-000C-4229
; Sequence 4229, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4229
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4229
```

```
Query Match 56.4%; Score 967; DB 4; Length 357;
Best Local Similarity 58.3%; Pred. No. 8e-89;
Matches 197; Conservative 44; Mismatches 91; Indels 6; Gaps 5;

QY 1 MVTGVGNGFGRIGRLAFRRIONV-EGVEVTRINDLTPNMLAHLKYDTTQGRFDGTV 59
Db 22 MVTGVGNGFGRIGRLAFRRIONV-EGVEVTRINDLTPNMLAHLKYDTTQGRFDGTV 81

QY 60 VKEGFEVNGNFIKVSARDPENIDWTD-GVEIVLEATGFFAKKEAAEKHLHANGAKV 118
Db 82 ATENGIVVDGEETRVYAEPAASKIPWVKENGVDIVLECTGFTYSEEKAAHLDA-GVKRV 140

QY 119 VITAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTI 178
Db 141 VISAPAG-AMKTIYVNVNDDTLNDKTIISAGSCTTNCLAPMAYFLNNEFGIEVGTMTTV 199

QY 179 HAYTGDMILDPHRRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPV 238
Db 200 HAYTSTQMLLDGPRVGGNLRRAARSADNTIPHSTGAAKAIGLVIPELNGKLDGAAQRPV 259

QY 239 PTGSVTELVTLKKNVSVWYDNEMSYTAQLVRLTLEYFAKIAK 334
Db 260 VDGSLTELVSILKTKVTADQVNEAMKHTIDNPSFGYDDREIVSGDIIGTTEGSIFDPTQ 319

QY 297 TKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 334
Db 320 TEVTTAGDFQLVKTVAWYDNEYGFTCMIRLLEKFAKIAK 357
```

```
RESULT 15
US-09-634-238-233
; Sequence 233, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
```

```
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-233

Query Match 56.3%; Score 966; DB 4; Length 340;
Best Local Similarity 57.0%; Pred. No. 9.4e-89;
Matches 195; Conservative 48; Mismatches 89; Indels 10; Gaps 6;

QY 1 MVTGVGNGFGRIGRLAFRRIONV-EGVEVTRINDLTPNMLAHLKYDTTQGRFDG 56
Db 1 MVTGVGNGFGRIGRLAFRRIONV-EGVEVTRINDLTPNMLAHLKYDTTQGRFDG 60

QY 57 TVEVKEGFEVNGNFIKVSARDPENIDWTD-GVEIVLEATGFFAKKEAAEKHLHANGA 115
Db 61 EVSATDNGIVVDGKEYRVYAEPAQNIWVKNDGVDYVLECTGFTYSEEKAAHLDA-GA 119

QY 116 KKVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLM 175
Db 120 KRVLISAPAGK-IKTIYVNVNDDTLNDKTIISAGSCTTNCLAPMAYFLNNEFGIEVGT 178

QY 176 TTIHAYTGDMILDPHRRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQ 235
Db 179 TTVHAYTSTQMLLDGPRVGGNLRRAARSADNTIPHSTGAAKAIGLVIPELNGKLDG 238

QY 236 VPVPTGSVTELVTLKKNVSVWYDNEMSYTAQLVRLTLEYFAKIAK 334
Db 239 VSVVDGSLTELVSILKTKVTADQVNEAMKHTIDNPSFGYDDREIVSGDIIGTTEGS 298

QY 293 DATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 334
Db 299 DPTQTEVTTAGDYQLVKTVAWYDNEYGFTCMIRLLEKFAKIAK 340

Search completed: May 11, 2004, 16:50:40
Job time : 42 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:45:46 ; Search time 21 Seconds
(without alignments)
1539.064 Million cell updates/sec

Title: US-10-650-369-12
Perfect score: 1715
Sequence: 1 MVVKVGINFGFRIGRLAFRR.....EMSYTAQLVRLTYFAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1711	99.8	336	2 A42963	glyceraldehyde-3-p
2	1617	94.3	336	2 S71350	glyceraldehyde-3-p
3	1593.5	92.9	335	2 F95235	glyceraldehyde 3-p
4	1593.5	92.9	359	2 G98099	glyceraldehyde-3-p
5	1410	82.2	336	2 F86905	hypothetical prote
6	1338.5	78.0	337	2 G86694	hypothetical prote
7	1289.5	75.2	334	2 S34254	glyceraldehyde-3-p
8	1262	73.6	336	2 AC1382	glyceraldehyde 3-p
9	1261	73.5	336	2 AD1751	glyceraldehyde 3-p
10	1259.5	73.4	334	2 C96987	glyceraldehyde 3-p
11	1197.5	69.8	334	2 B82019	glyceraldehyde-3-p
12	1184.5	69.1	334	2 E81001	glyceraldehyde 3-p
13	1183	69.0	336	2 E89850	glyceraldehyde-3-p
14	1003	58.5	338	2 T09633	glyceraldehyde-3-p
15	993	57.9	333	2 F90881	glyceraldehyde 3-p
16	993	57.9	333	2 C85737	glyceraldehyde-3-p
17	921	53.7	337	2 S73737	glyceraldehyde-3-p
18	909.5	53.0	336	2 A43260	glyceraldehyde-3-p
19	908	52.9	337	2 C64233	glyceraldehyde-3-p
20	907	52.9	349	2 F90517	glyceraldehyde 3-p
21	896.5	52.3	335	1 DEBSG	glyceraldehyde-3-p
22	891.5	52.0	335	2 S12696	glyceraldehyde-3-p
23	885.5	51.6	335	1 DEBSGF	glyceraldehyde-3-p
24	882.5	51.5	335	2 H84094	glyceraldehyde-3-p
25	874	51.0	333	1 DEHGGT	glyceraldehyde-3-p
26	863	50.3	335	2 A70107	probable glyceral
27	843	49.2	342	2 F70391	glyceraldehyde-3-p
28	842.5	49.1	336	2 T36020	glyceraldehyde-3-p
29	831	48.5	334	2 AI0262	glyceraldehyde-3-p

30	830	48.4	331	2 F82131	glyceraldehyde 3-p
31	823	48.0	330	2 E75408	glyceraldehyde 3-p
32	822	47.9	331	2 AG0711	glyceraldehyde 3-p
33	818	47.7	331	1 DEECG3	glyceraldehyde-3-p
34	818	47.7	331	2 H90939	glyceraldehyde-3-p
35	818	47.7	331	2 D85788	glyceraldehyde-3-p
36	816	47.6	339	2 G70915	glyceraldehyde-3-p
37	815	47.5	339	2 S72763	glyceraldehyde-3-p
38	807.5	47.1	339	2 G64041	glyceraldehyde-3-p
39	799	46.6	331	1 DEUTC	glyceraldehyde-3-p
40	787.5	45.9	336	2 T40235	glyceraldehyde-3-p
41	781	45.5	337	2 S54141	glyceraldehyde-3-p
42	779.5	45.5	331	1 DETWG3	glyceraldehyde-3-p
43	779.5	45.5	337	2 S40610	glyceraldehyde-3-p
44	777.5	45.3	332	2 JC4373	glyceraldehyde-3-p
45	775	45.2	344	2 E84043	glyceraldehyde-3-p

ALIGNMENTS

RESULT 1

A42963
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus
N;Alternate names: plasmin receptor
C;Species: Streptococcus sp.
C;Date: 10-Mar-1994 #sequence revision 10-Mar-1994 #text_change 03-Jun-2002
C;Accession: A42963; B42963; JH0750
R;Lottenberg, R.; Broder, C.C.; Boyle, M.D.; Kain, S.J.; Schroeder, B.L.; Curtiss III, F
J. Bacteriol. 174, 5204-5210, 1992
A;Title: Cloning, sequence analysis, and expression in Escherichia coli of a streptococ
A;Reference number: A42963; MUID:92355491; PMID:1322883
A;Accession: A42963
A;Molecule type: DNA
A;Residues: 1-336 <LOT>
A;Experimental source: group A, strain 64/14
A;Note: sequence extracted from NCBI backbone (NCBIP:110308)
A;Accession: B42963
A;Molecule type: protein
A;Residues: 2-74,161-164,'X',166-174;187-211,'X',213-217 <LO2>
R;Pancholi, V.; Fischetti, V.A.
J. Exp. Med. 176, 415-426, 1992
A;Title: A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate
A;Reference number: JH0750; MUID:92364544; PMID:1500854
A;Accession: JH0750
A;Molecule type: protein
A;Residues: 2-30,'A',32-40 <PAN>
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F;152/Active site: Cys #status predicted

Query Match 99.8%; Score 1711; DB 2; Length 336;
Best Local Similarity 99.4%; Pred. No. 4.2e-116;
Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MVVKVGINFGFRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKDYDTTQGRFDGTVEV	60
Db	1	MVVKVGINFGFRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKDYDTTQGRFDGTVEV	60
Qy	61	KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI	120
Db	61	KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI	120
Qy	121	TAPGGNDVKTVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA	180
Db	121	TAPGGNDVKTVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA	180
Qy	181	YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT	240
Db	181	YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT	240
Qy	241	GSVTELVTLDKNVSYDEINAAKASNDISFGYTEDPIVSSDIVGSYSLFDATQTQVM	300
Db	241	GSVTELVTLDKNVSYDEINAAKASNDISFGYTEDPIVSSDIVGSYSLFDATQTQVM	300

Db 241 GSVTELVVTLTKNSVDENSAKKAASNDSTGTYEDPIVSSDIVGVSGLFDATQTKVM 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 2

S71350

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus

C:Species: Streptococcus "equisimilis"

C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 03-Jun-2002

C:Accession: S71350

R:Gase, K.; Gase, A.; Schirmer, H.; Malke, H.

Eur. J. Biochem. 239, 42-51, 1996

A:Title: Cloning, sequencing and functional overexpression of the Streptococcus equisimilis

inding protein. Purification and biochemical characterization of the protein.

A:Reference number: S71350; MUID:96305364; PMID:8706717

A:Accession: S71350

A:Molecule type: DNA

A:Residues: 1-336 <GAS>

A:Cross-references: EMBL:X97788; NID:g1478268; PIDN:CAA66377.1; PID:g1478269

A:Experimental source: strain H46A

C:Genetics:

A:Gene: gapC

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: gluconeogenesis; glycolysis; NAD; oxidoreductase

F:4-34/Region: beta-alpha-beta NAD nucleotide-binding fold

F:152/Active site: Cys #status predicted

Query Match 94.3%; Score 1617; DB 2; Length 336;

Best Local Similarity 95.5%; Pred. No. 2.5e-109;

Matches 321; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 60

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 60

Qy 61 KEGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120

Db 61 KEGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120

Qy 121 TAPGNDVKTIVFNTNHDILDTGTTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180

Db 121 TAPGNDVKTIVFNTNHDILDTGTTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180

Qy 181 YTGDMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Db 181 YTGDMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Qy 241 GSVTELVVTLTKNSVDENSAKKAASNDSTGTYEDPIVSSDIVGVSGLFDATQTKVM 300

Db 241 GSVTELVVTLTKNSVDENSAKKAASNDSTGTYEDPIVSSDIVGVSGLFDATQTKVM 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 3

F95235

glyceraldehyde 3-phosphate dehydrogenase [imported] - Streptococcus pneumoniae (strain T

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C:Accession: F95235

R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: F95235

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-335 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK76079.1; PID:g14973522; GSPDB:GN00164; TIGR.SP4

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP2012

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 92.9%; Score 1593.5; DB 2; Length 335;

Best Local Similarity 92.6%; Pred. No. 1.3e-107;

Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

Qy 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 60

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 60

Qy 61 KEGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120

Db 61 KEGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 119

Qy 121 TAPGNDVKTIVFNTNHDILDTGTTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180

Db 121 TAPGNDVKTIVFNTNHDILDTGTTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 179

Qy 181 YTGDMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

Db 181 YTGDMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 239

Qy 241 GSVTELVVTLTKNSVDENSAKKAASNDSTGTYEDPIVSSDIVGVSGLFDATQTKVM 300

Db 241 GSVTELVVTLTKNSVDENSAKKAASNDSTGTYEDPIVSSDIVGVSGLFDATQTKVM 299

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 4

G98099

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [imported] - St

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002

C:Accession: G98099

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; F

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; N

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: G98099

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-359 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAL00628.1; PID:g15459513; GSPDB:GN00174

C:Genetics:

A:Gene: gapA

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 92.9%; Score 1593.5; DB 2; Length 359;

Best Local Similarity 92.6%; Pred. No. 1.4e-107;

Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

Qy 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 60

Db 25 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 84

Qy 61 KEGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120

Db 85 KEGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 143

QY 121 TAPGGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 144 TAPGGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALQDNFVVEGLMTTIHA 203
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 204 YTGDMILDGPHRGDLRRARAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGSAQRPVPT 263
QY 241 GSVTELVVTLDKNVSVDENAAKAAASNDSPGYTEDPIVSSDIVGSYSLFDTATQTKVM 300
Db 264 GSVTELVVTLDKNVSVDENAAKAAASNDSPGYTEDPIVSSDIVGSYSLFDTATQTKVL 323
QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 324 DVDGKQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 359

RESULT 5
F86905
hypothetical protein gapB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: F86905
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: F86905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-336 <STO>
A;Cross-references: GB:AE005176; PID:g12725315; PIDN:AAK06344.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: gapB
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 82.2%; Score 1410; DB 2; Length 336;
Best Local Similarity 81.2%; Pred. No. 2.2e-94;
Matches 273; Conservative 22; Mismatches 41; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRILAFRRIONVEGVEVTRINDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRILAFRRIONVEGVEVVAINDLTPAMLAHLKLYDTTQGRFDGKVEV 60
QY 61 KEGGFEVNGNFIKVSABRDPENIDWATDGVLEIVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGKFKVTAESNPANINWAEVGAIEIVLEATGFFATKEAEQHLHANGAKKVI 120
QY 121 TAPGGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGSDVKTIVFNTHDILDGTETVISGASCTTNCLAPMADTLNKQFGIKVGTMTTVHG 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMTLDGPHRGDLRRARAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGHAQRPVPT 240
QY 241 GSVTELVVTLDKNVSVDENAAKAAASNDSPGYTEDPIVSSDIVGSYSLFDTATQTKVM 300
Db 241 GSVTELVVTLDKNVSVDENAAKAAASNDSPGYTEDPIVSSDIVGINSLSLFDATQTEVT 300
QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 SADGAQLVKVSVWYDNEMSYTSLNLRVTLAYFAKIAK 336

RESULT 6
G86694
hypothetical protein gapA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: G86694
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: G86694
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-337 <STO>
A;Cross-references: GB:AE005176; PID:g12723446; PIDN:AAK04657.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: gapA
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 78.0%; Score 1338.5; DB 2; Length 337;
Best Local Similarity 77.2%; Pred. No. 3.2e-89;
Matches 260; Conservative 31; Mismatches 45; Indels 1; Gaps 1;

QY 1 MVVKVINGFGRIGRILAFRRIONVEGVEVTRINDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRILAFRRIONVEGVEVAHINDLTPAMLAHLKLYDTTQGRFKGTVEV 60
QY 61 KEGGFEVNGNFIKVSABRDPENIDWATDGVLEIVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGDFVNGKFKVTAERNPEDIQWADSGVEIVLEATGFFATKEAEKHLHPGGAKKVI 120
QY 121 TAPGGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGNDVKT VVFNTHDILDGTETVISAGSCITNSLAPMADALNKNFGVKGSTMTVHS 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMTLDGPHRGDLRRARAGAAANIVPNSGTGAAGAIGLVIPELSGLMKGHQAVSTPT 240
QY 241 GSVTELVVTLDKNVSVDENAAKAAASNDSPGYTEDPIVSSDIVGSYSLFDTATQTKVM 300
Db 241 GSITELVTVLEKHTVDEINEAMKAAADESFGVNVDEIVSSDIIGMAYGSLFDTATLTVT 300
QY 301 EV-DGSQVLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 DLKGGQLVKTAAWYDNEMSFATQILRTLEYFAKIAK 337

RESULT 7
S34254
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Clostridium
C;Species: Clostridium pasteurianum
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002
C;Accession: S34254
R;Oster, T.; Assobhei, O.; Scherrer, S.; Branlant, G.; Branlant, C.
submitted to the EMBL Data Library, May 1993
A;Description: Nucleotide sequence of the glyceraldehyde-3-phosphate dehydrogenase from
A;Reference number: S34254
A;Accession: S34254
A;Molecule type: DNA
A;Residues: 1-334 <OST>
A;Cross-references: EMBL:X7219; NID:g311923; PIDN:CAA51020.1; PID:g311924
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: oxidoreductase

Query Match 75.2%; Score 1289.5; DB 2; Length 334;
Best Local Similarity 74.3%; Pred. No. 1.1e-85;
Matches 249; Conservative 33; Mismatches 52; Indels 1; Gaps 1;

QY 2 VVKVINGFGRIGRILAFRRIONVEGVEVTRINDLTPNMLAHLKLYDTTQGRFDGTVEV 61
Db 1 MTKVAINGFGRIGRILAFRRILEVPGLEVVAINDLTDKMLAHLFKYDSSQGRFNGEIEVK 60
QY 62 EGGFEVNGNFIKVSABRDPENIDWATDGVLEIVLEATGFFAKKEAEKHLHANGAKKVI 121
Db 61 EGAFVNGKFKVFAEADPEKLPWGELGIDVVLVLECTGFFTKKEAEHVRA-GAKKVIS 119
QY 122 APGGNDVKT VVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 181

Db 120 APAGNDLKTIVFNVNNEEDLDGTETVISGASCTTNCCLAPMAKVLNDFKFGIEKGFMTTIHAY 179

QY 182 TGDQMILDPGHRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTG 241

Db 180 TNDQNTLDGPHRGKDFRRARAAVSIIPNSTGAAKAIAQVIPELKGLDGNAGRPVPTG 239

QY 242 SVTELVTLDKNVSVDEINAAKMAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKVME 301

Db 240 SVTELLSVLKKNVTVVEEINAAKMAANESFGYTEDEIVSADWVGISYGLSFDATLTKIVD 299

QY 302 VDGSQLVKVSWYDNEMSYTAQLVRTLLEYFAKIAK 336

Db 300 VDGSQLVKVSWYDNEMSYTSQLVRTLLEYFAKIAK 334

RESULT 8

AC1382

glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria monocytogenes

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AC1382

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, Science 294, 849-852, 2001

A:Title: Comparative Genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1382

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-336 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAD00537.1; PID:g16411947; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: gap

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.6%; Score 1262; DB 2; Length 336;

Best Local Similarity 73.3%; Pred. No. 1.1e-83;

Matches 247; Conservative 33; Mismatches 55; Indels 2; Gaps 2;

QY 1 MVTVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKLYDTTQGRFDGVEV 60

Db 1 MVTVKVINGFGRIGRLAFRRIONVEGVEVAINDLTDKMLAHLKLYDTTQGRFDGVEV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120

Db 61 HDGFFKNGKEVKVLARNRNPGLDGLVDIVLECTGFFTAQDKAELHIKA-GAKKVI 119

QY 121 TAPGGNDVKTIVFNTNHDILDTETVISGASCTTNCCLAPMAKALHDAFGIQKGLMTTIHA 180

Db 120 SAPATGDMKTIVYNNVNHETLDGTETVISGASCTTNCCLAPMAKALHDAFGIQKGLMTTIHA 179

QY 181 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

Db 180 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 239

QY 241 GSVTELVTLTKNVSVDEINAAKMAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKV 299

Db 240 GSVTELVTLTKNVSVDEINAAKMAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKV 299

QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRTLLEYFAKIAK 336

Db 300 LTVGQQQLVKTVAWYDNEMSYTAQLVRTLLEYFAKIAK 336

RESULT 9

AD1751

glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria innocua (strain AD1751)

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AD1751

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, Science 294, 849-852, 2001

A:Title: Comparative Genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AD1751

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-336 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97780.1; PID:g16415075; GSPDB:GN00178

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: gap

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.5%; Score 1261; DB 2; Length 336;

Best Local Similarity 73.3%; Pred. No. 1.3e-83;

Matches 247; Conservative 32; Mismatches 56; Indels 2; Gaps 2;

QY 1 MVTVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKLYDTTQGRFDGVEV 60

Db 1 MVTVKVINGFGRIGRLAFRRIONVEGVEVAINDLTDKMLAHLKLYDTTQGRFDGVEV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120

Db 61 HDGFFVNGKEVKVLARNRNPGLDGLVDIVLECTGFFTAQDKAELHIKA-GAKKVI 119

QY 121 TAPGGNDVKTIVFNTNHDILDTETVISGASCTTNCCLAPMAKALHDAFGIQKGLMTTIHA 180

Db 120 SAPATGDMKTIVYNNVNHETLDGTETVISGASCTTNCCLAPMAKALHDAFGIQKGLMTTIHA 179

QY 181 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

Db 180 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 239

QY 241 GSVTELVTLTKNVSVDEINAAKMAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKV 299

Db 240 GSVTELVTLTKNVSVDEINAAKMAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKV 299

QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRTLLEYFAKIAK 336

Db 300 LTVGQQQLVKTVAWYDNEMSYTAQLVRTLLEYFAKIAK 336

RESULT 10

C96987

glyceraldehyde 3-phosphate dehydrogenase, gene gapC [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: C96987

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C96987

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK78686.1; PID:g15023589; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0709

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.4%; Score 1259.5; DB 2; Length 334;

Best Local Similarity 71.9%; Pred. No. 1.6e-83;

Matches 241; Conservative 40; Mismatches 53; Indels 1; Gaps 1;

C;Genetics:
A;Gene: gap
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 69.0%; Score 1183; DB 2; Length 336;
Best Local Similarity 68.2%; Pred. No. 5.3e-78;
Matches 230; Conservative 43; Mismatches 62; Indels 2; Gaps 2;

QY 1 MVVKGINGFGRIGRLAFRRIQNVGVEVTRINDLTDPNMLAHLKDYDTQGRFDGTVEV 60
Db 1 MAVKVAINGFGRIGRLAFRRIQVEGLEVVAVNDLTDMLAHLKDYDTMQGRFTGEVEV 60

QY 61 KEGGFEVNGFIKVSARDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVVI 120
Db 61 VDGGRVNGKEVKSFPSEPDASKLPWKDLNIDVLECTGYFTDKOKAQAHIEA-GAKKVL 119

QY 121 TAPGGNDVKTIVFNTNHDLDGTETVISGASCTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 120 SAPATGDLKTIVFNTNHDLDGTETVISGASCTNCLAPMAKALHDAFGIQKGLMTTIHA 179

QY 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 180 YTGDMQTDAPHRKGRARARAAENIIPNSTGAAGAIGKVIPEIDGKLDGGAQRPVPT 239

QY 241 GSVTELVTLDK-NVSVDEINAAMKAASNDSPFGYTEDPIVSSDIVGVSGLFDATQTKV 299
Db 240 GSVTELVTLDK-NVSVDEINAAMKAASNDSPFGYTEDPIVSSDIVGVSGLFDATQTKV 299

QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 300 MSVGDRLVKVAAWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 14
T09633
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Lactobacillus
C;Species: Lactobacillus delbrueckii
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C;Accession: T09633

R;Branny, P.; Delatorre, F.; Gare, J.R.
Microbiology 144, 905-914, 1998
A;Title: An operon coding for three glycolytic enzymes in Lactobacillus delbrueckii subsp.

A;Reference number: Z16788; MUID:98240227; PMID:9579064
A;Accession: T09633
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-338 <BRA>
A;Cross-references: EMBL:AJ000339; NID:G2624189; PIDN:CAA04014.1; PID:G2624191
A;Experimental source: subsp. bulgaricus, strain B107
C;Genetics:

A;Gene: gap
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match 58.5%; Score 1003; DB 2; Length 338;
Best Local Similarity 58.7%; Pred. No. 5.2e-65;
Matches 199; Conservative 45; Mismatches 89; Indels 6; Gaps 3;

QY 1 MVVKGINGFGRIGRLAFRRI---QNVGVEVTRINDLTDPNMLAHLKDYDTQGRFDG 56
Db 1 MTVKIGINGFGRIGRLAFRRIMDLGEETKDIEVVAINDLTTPAMLAHLKDYDSTHGTFDH 60

QY 57 TVEVKEGGFEVNGNFIKVSARDPENIDWA-TDGVVEIVLEATGFFAKKEAAEKHLHANGA 115
Db 61 EVSATEDSLVVDGKKYRVYAEPPQAQNPWVKNDGVDFVLECTGYFTSKAKSQAHLDA-GA 119

QY 116 KKVVITAPGGNDVKTIVFNTNHDLDGTETVISGASCTNCLAPMAKALHDAFGIQKGLM 175
Db 120 KRVLISAPAGNDLKTIVSVNQDTLTADDTIVSAGSCTNSLAPMANALNKEFGIQVGTM 179

QY 176 TTIHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQR 235

Db 180 TTIHAYTATQKVLDPDRGNFRNARAAENIIPHSTGAAGAIGLVLPELNGKLDGHAQR 239
QY 236 VPVPTGTSVTELVTLDKNVSVDEINAAMKAASNDSPFGYTEDPIVSSDIVGVSGLFDAT 295
Db 240 VPVKDGSSETELVTILDKKVTAEVNAAMKKYSPSFAYNADQIVSTDVLGWTAGSIFDPT 299
QY 296 QTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAK 334
Db 300 QTQVITAGDKQLVKTVAWYDNEYSFTCQMVRLTLHFATL 338

RESULT 15
F90881
glyceraldehyde 3-phosphate dehydrogenase C [imported] - Escherichia coli (strain O157:H
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: F90881
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90881
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA035445.1; PID:G13361488; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2022
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 57.9%; Score 993; DB 2; Length 333;
Best Local Similarity 59.3%; Pred. No. 2.7e-64;
Matches 198; Conservative 54; Mismatches 76; Indels 6; Gaps 5;

QY 4 KVGINGFGRIGRLAFRRIQNVGVEVTRINDLTDPNMLAHLKDYDTQGRFDGTVEVKE 62
Db 3 KVGINGFGRIGRLVLRLLLEKSNIDVAINDLTSPKILAYLLKHDSNYGPPFWSVDFTE 62

QY 63 GGFVNGNFIKVSARDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVITA 122
Db 63 DSLIVDGKSIAYAEKEAKNIPWKAKGAEIIVECTGYFTSAEKSOAHLDA-GAKKVLISA 121

QY 123 PGNDVKTIVFNTNHDLDGTETVISGASCTNCLAPMAKALHDAFGIQKGLMTTIHAYT 182
Db 122 PAG-EMKTIIVYKVNDDTLTGNDTIVSVASCTNCLAPMAKALHDSFGIEVGTMTTIHAYT 180

QY 183 GDMILDGPHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGS 242
Db 181 GTQSLVDGP-RGKOLRASRAAAENIIPHTTGAAGAIGLVIPELNGKLDGAAQRPVPTGS 239

QY 243 VTELVTLDKNVSVDEINAAMKAA--SNDSPFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 240 VTELVSILGKKVTAEVNAALNQATNNESFGYTDDEIVSSDIIGSHFGSVDFDQTQTEIT 299

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAK 334
Db 300 AVGDLQLVKTVAWYDNEYGFVTLIRTLEKFAKL 333

Search completed: May 11, 2004, 16:49:49
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:42:06 ; Search time 18 Seconds
(without alignments)
971.976 Million cell updates/sec

Title: US-10-650-369-12

Perfect score: 1715

Sequence: 1 MVVKVINGFGRIQLAFRR.....EMSYTAQLVRLTYFAKIAK 336

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1709	99.7	335	1 G3P_STRPY	P50467 streptococ
2	1705	99.4	335	1 G3P_STRP3	Q8K8m9 streptococ
3	1612	94.0	335	1 G3P_STRPQ	Q59906 streptococ
4	1338.5	78.0	337	1 G3P_LACLA	P52987 lactococcc
5	1289.5	75.2	334	1 G3P_CLOPA	Q59309 clostridium
6	1259.5	73.4	334	1 G3P_CLOAB	O52631 clostridium
7	1190	69.4	336	1 G3P1_STAEP	Q8cyp5 staphylococ
8	1183	69.0	336	1 G3P1_STAAM	Q925c5 staphylococ
9	1003	58.5	338	1 G3P_LACDE	O32755 lactobacill
10	993	57.9	333	1 G3P3_ECO57	P58072 escherichia
11	992	57.8	333	1 G3P3_ECOLI	P33898 escherichia
12	921	53.7	337	1 G3P_MYCPN	P75358 mycoplasma
13	908.5	53.0	334	1 G3P_CORGL	Q01651 corynebacte
14	908	52.9	337	1 G3P_MYCGE	P47543 mycoplasma
15	891.5	52.0	334	1 G3P1_BACSU	P09124 bacillus su
16	886.5	51.7	334	1 G3P_BACME	P23722 bacillus me
17	880.5	51.3	334	1 G3P_BACST	P00362 bacillus st
18	873	50.9	332	1 G3P_THEMEA	P17721 thermotoga
19	863	50.3	335	1 G3P_BORBU	P46795 borrelia bu
20	843	49.2	342	1 G3P_AQUAE	O67161 aquifex aeo
21	842.5	49.1	336	1 G3P_STRCO	Q92518 streptomyce
22	817	47.6	330	1 G3P1_SALTY	P24165 salmonella
23	816	47.6	339	1 G3P_MYCTU	O06822 mycobacteri
24	815	47.5	339	1 G3P_MYCLE	P46713 mycobacteri
25	813	47.4	330	1 G3P1_ECOLI	P06977 escherichia
26	807.5	47.1	339	1 G3P_HAEIN	P44304 haemophilus
27	801	46.7	339	1 G3P_MYCAV	P94915 mycobacteri
28	794.5	46.3	337	1 G3P1_RHIRA	Q9c136 rhizomucor
29	794	46.3	330	1 G3PC_TRYBB	P10097 trypanosoma
30	787.5	45.9	336	1 G3P1_SCHPO	P78958 schizosacch
31	784.5	45.7	332	1 G3P_RALSO	P52694 ralstonia s
32	783.5	45.7	337	1 G3P_MONAN	P53430 monascus an
33	779.5	45.5	331	1 G3P_THEAQ	P00361 thermus aqu

34	779.5	45.5	337	1	G3P_CLAPU	Q00584 claviceps p
35	777.5	45.3	332	1	G3P_STRAU	Q59800 streptomyce
36	775	45.2	336	1	G3P2_SYNY3	P80505 synechocyst
37	772	45.0	333	1	G3P_STRAE	P54226 streptomyce
38	772	45.0	337	1	G3P2_ANASP	P58554 anabaena sp
39	771.5	45.0	337	1	G3P_COCHE	P29497 cochliobolu
40	771	45.0	330	1	G3P2_LEIME	Q01558 leishmania
41	770.5	44.9	335	1	G3P2_SCHPO	O43026 schizosacch
42	770.5	44.9	337	1	G3P_CURLU	P28844 curvularia
43	768.5	44.8	336	1	G3P_ASPNG	Q12552 aspergillus
44	766.5	44.7	336	1	G3P_EMENI	P20445 emericella
45	762.5	44.5	338	1	G3P_NEUCR	P54118 neurospora

ALIGNMENTS

RESULT 1

G3P_STRPY	STANDARD;	PRT;	335 AA.
AC P50467;			
DT 01-OCT-1996 (Rel. 34, Created)			
DT 16-OCT-2001 (Rel. 40, Last sequence update)			
DT 28-FEB-2003 (Rel. 41, Last annotation update)			
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)			
DE (Plasminogen-binding protein) (Plasmin receptor).			
GN GAP OR PLR OR GAPA OR SPY0274 OR SPYM18_0261.			
OS Streptococcus pyogenes, and			
OS Streptococcus pyogenes (serotype M18).			
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC Streptococcus.			
OX NCBI_TaxID=1314, 186103;			
RN [1]			
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-73; 160-173 AND 186-216.			
RC STRAIN=64/14;			
RX MEDLINE=92355491; PubMed=1322883;			
RA Lottenberg R., Broder C.C., Boyle M.D., Kain S.J., Schroeder B.L.,			
RA Curtiss R. III;			
RT "Cloning, sequence analysis, and expression in Escherichia coli of a			
RT streptococcal plasmin receptor."			
RL J. Bacteriol. 174:5204-5210(1992).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;			
RX MEDLINE=21192684; PubMed=11296296;			
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,			
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,			
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,			
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;			
RT "Complete genome sequence of an M1 strain of Streptococcus			
RT pyogenes."			
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).			
RN [3]			
RP SEQUENCE FROM N.A.			
RC STRAIN=MGAS8232 / Serotype M18;			
RX MEDLINE=21927593; PubMed=11917108;			
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,			
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,			
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,			
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;			
RT "Genome sequence and comparative microarray analysis of serotype M18			
RT group A Streptococcus strains associated with acute rheumatic fever			
RT outbreaks."			
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).			
RN [4]			
RP SEQUENCE OF 1-11; 20-30; 103-128; 162-171 AND 199-215.			
RC STRAIN=JRS4 / Serotype M6;			
RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,			
RA VanBogelen R.A.;			
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes			
RT proteins."			
RL Submitted (MAY-2000) to Swiss-Prot.			
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.			

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M95569; AAA26953.1; -.
CC EMBL; AE006494; AAK33348.1; -.
CC EMBL; AE009973; AAL97041.1; -.
CC HSSP; P00362; 1GD1.
CC InterPro; IPR000173; GAP dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
KW
FT INIT MET 0 0
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
FT CONFLICT 261 261 A -> S (IN REF. 1).
FT SEQUENCE 335 AA; 35811 MW; F06006EE253C8A3F CRC64;
SQ
Query Match 99.7%; Score 1709; DB 1; Length 335;
Best Local Similarity 99.7%; Pred. No. 3.4e-111;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 VVKVINGFGRIGRLAFRIQNVGEVETRLNDLTPNMLAHLKYYDTTQGRFDGTVEVK 61
DB 1 VVKVINGFGRIGRLAFRIQNVGEVETRLNDLTPNMLAHLKYYDTTQGRFDGTVEVK 60
QY 62 EGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVIT 121
DB 61 EGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVIT 120
QY 122 APGGNDVKTTFVNTNHDILDTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 181
DB 121 APGGNDVKTTFVNTNHDILDTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 180
QY 182 TGDQMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTG 241
DB 181 TGDQMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTG 240
QY 242 SVTELVTLDKNVSVDEINAAKAAASNDSPGYTEDPIVSSDIVGSYGLFDPATQTKVME 301
DB 241 SVTELVTLDKNVSVDEINAAKAAASNDSPGYTEDPIVSSDIVGSYGLFDPATQTKVME 300
QY 302 VDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 336
DB 301 VDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 335
RESULT 2
G3P_STRP3 STANDARD; PRT; 335 AA.
ID G3P_STRP3
AC Q8K8M9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)

DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR PLR OR SPYM3_0201 OR SPS0207.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RA "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=SI-1 / Serotype M3;
RC MEDLINE=22683278; PubMed=12799345;
RX Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RA "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RT Genome Res. 13:1042-1055(2003).
RL [1]
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN (By similarity).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014140; AAM78808.1; -.
CC EMBL; AP005141; BAC63302.1; -.
CC InterPro; IPR000173; GAP dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
KW
FT INIT MET 0 0
FT BINDING 151 151 BY SIMILARITY.
FT ACT_SITE 178 178 GLYCERALDEHYDE 3-PHOSPHATE (BY
FT SEQUENCE 335 AA; 35841 MW; 4DCB76F382F6F98 CRC64;
SQ
Query Match 99.4%; Score 1705; DB 1; Length 335;
Best Local Similarity 99.4%; Pred. No. 6.3e-111;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 VVKVINGFGRIGRLAFRIQNVGEVETRLNDLTPNMLAHLKYYDTTQGRFDGTVEVK 61
DB 1 VVKVINGFGRIGRLAFRIQNVGEVETRLNDLTPNMLAHLKYYDTTQGRFDGTVEVK 60
QY 62 EGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVIT 121

Db 61 EGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHTNGAKKVVIT 120

QY 122 APGGNDVKTVVNTNHDILDTGTTVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 181

Db 121 APGGNDVKTVVNTNHDILDTGTTVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 180

QY 182 TGDQMILDGPHRGDLRRARAGAANIVPNTGAAKAIGLVIPELNGKLDGAAQRPVPTG 241

Db 181 TGDQMILDGPHRGDLRRARAGAANIVPNTGAAKAIGLVIPELNGKLDGAAQRPVPTG 240

QY 242 SVTELVTILDKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKVME 301

Db 241 SVTELVTILDKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKVME 300

QY 302 VDGSQLVKVSVDNEMSYTAQLVRLTLEYFAKIAK 336

Db 301 VDGSQLVKVSVDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 3

G3P STREQ

ID - G3P STREQ STANDARD; PRT; 335 AA.

AC Q59906;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)

DE (plasminogen-binding protein) (plasmin receptor).

GN GAP OR GAPC.

OS Streptococcus equisimilis.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=119602;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H46A;

RX MEDLINE=96305364; PubMed=8706717;

RA Gase K., Gase A., Schirmer H., Malke H.;

RT "Cloning, sequencing and functional overexpression of the

RT Streptococcus equisimilis H46A gapC gene encoding a

RT glyceraldehyde-3-phosphate dehydrogenase that also functions as a

RT plasmin(ogen)-binding protein. Purification and biochemical

RT characterization of the protein.";

RL Eur. J. Biochem. 239:42-51(1996).

CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +

CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -!- PATHWAY: Second phase of glycolysis; first step.

CC -!- SUBUNIT: Homotetramer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate

CC dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

EMBL; X97788; CAA66377.1; -.

EMBL; Y12602; CAA73174.1; -.

PIR; S71350; S71350.

HSSP; P00362; 1GD1.

InterPro; IPR000173; GAP dhdrogenase.

InterPro; IPR006424; GAPDH-I.

Pfam; PF00044; gpdh; 1.

Pfam; PF02800; gpdh C; 1.

PRINTS; PR00078; G3PDHGRNASE.

TIGRFAMs; TIGR01534; GAPDH-I; 1.

PROSITE; PS00071; GAPDH; 1.

KW Glycolysis; Oxidoreductase; NAD.

FT INIT MET 0 0 BY SIMILARITY.

FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY

FT ACT_SITE 178 178 SIMILARITY).

FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS

FT SEQUENCE 335 AA; 35739 MW; FE7ACFD7663E46 CRC64;

SQ

Query Match 94.0%; Score 1612; DB 1; Length 335;

Best Local Similarity 95.5%; Pred. No. 1.7e-104;

Matches 320; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 2 VVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTDPNMLAHLKYDTTQGRFDGTVEVK 61

Db 1 VVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTDPNMLAHLKYDTTQGRFDGTVEVK 60

QY 62 EGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVVIT 121

Db 61 EGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVVIT 120

QY 122 APGGNDVKTVVNTNHDILDTGTTVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 181

Db 121 APGGNDVKQLFSTLTTISLDGTTVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 180

QY 182 TGDQMILDGPHRGDLRRARAGAANIVPNTGAAKAIGLVIPELNGKLDGAAQRPVPTG 241

Db 181 TGDQMIVDGHRRGGDLRRARAGAANIVPNTGAAKAIGLVIPELNGKLDGAAQRPVPTG 240

QY 242 SVTELVTILDKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKVME 301

Db 241 SVTELVTILDKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKVME 300

QY 302 VDGSQLVKVSVDNEMSYTAQLVRLTLEYFAKIAK 336

Db 301 VDGSQLVKVSVDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 4

G3P_LACLA

ID - G3P LACLA STANDARD; PRT; 337 AA.

AC P52987;

DT 01-OCT-1996 (Rel. 34, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).

GN GAP OR LL0559.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

OX NCBI_TaxID=1360;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LM0230;

RX MEDLINE=95291425; PubMed=7773380;

RA Cancilla M.R., Hillier A.J., Davidson B.E.;

RT "Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene,

RT gap: further evidence for strongly biased codon usage in glycolytic

RT pathway genes.";

RL Microbiology 141:1027-1036(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=IL1403;

RX MEDLINE=21235186; PubMed=11337471;

RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarre K.,

RA Weissbach J., Ehrlich S.D., Sorokin A.;

RT "The complete genome sequence of the lactic acid bacterium Lactococcus

RT lactis ssp. lactis IL1403.";

RL Genome Res. 11:731-753(2001).

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +

CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -!- PATHWAY: Second phase of glycolysis; first step.

CC -!- SUBUNIT: Homotetramer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate

CC dehydrogenase family.

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L36907; AAC41453.1; -.
CC EMBL; AE006290; AAK04657.1; -.
CC PIR; G86694; G86694.
CC HSSP; P17721; 1HDG.
CC InterPro; IPR000173; GAP dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; FALSE NEG.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT BINDING 152 152 GLYCERALDEHYDE 3-PHOSPHATE
FT (BY SIMILARITY).
FT ACT_SITE 179 179 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
FT CONFLICT 143 143 T -> S (IN REF. 1).
FT SEQUENCE 337 AA; 36057 MW; 17BB8C6AAEFF589D CRC64;
CC -----
Query Match 78.0%; Score 1338.5; DB 1; Length 337;
Best Local Similarity 77.2%; Pred. No. 1.4e-85;
Matches 260; Conservative 31; Mismatches 45; Indels 1; Gaps 1;
CC 1 MVVKVINGFGRIGRLAFRRIONVEGEVETRLNDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
CC 1 MVVKVINGFGRIGRLALRRIOVEGEVETRLNDLTDPNMLAHLKDYDTTQGRFKEV 60
CC 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
CC 61 KEGGFDVNGKFKVTAERNPEDIQWADSGVEIVLEATGFFATKEAEKHLHPPGAKKVI 120
CC 121 TAPGNDVKTIVFNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIIA 180
CC 121 TAPGNDVKTIVFNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIIA 180
CC 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
CC 181 YTGDMILDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
CC 241 GSVTELVTLTKNVSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
CC 241 GSVTELVTLTKNVSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
CC 301 EV-DGSQLVKVSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 336
CC 301 EV-DGSQLVKVSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 336
CC 301 DLKGGQLVKTAAYDNEMSFTAQLRTLEYFAKIAK 337
CC -----
RESULT 5
G3P_CLOPA STANDARD; PRT; 334 AA.
ID G3P_CLOPA
AC Q59309;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) (CP
DE 17/CP 18).
GN GAP.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Oster T., Assobhei O., Scherrer S., Branlant G., Branlant C.;
RA Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PARTIAL SEQUENCE OF 1-26.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RL sequence analysis of proteins from Clostridium pasteurianum W5.";
CC Electrophoresis 19:802-806(1998).
CC -|- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -|- PATHWAY: Second phase of glycolysis; first step.
CC -|- SUBUNIT: Homotetramer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X72219; CAA51020.1; -.
CC PIR; S34254; S34254.
CC HSSP; P00362; 1GD1.
CC InterPro; IPR000173; GAP dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD.
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
FT SEQUENCE 334 AA; 36078 MW; D15905D0DA7F62E7 CRC64;
CC -----
Query Match 75.2%; Score 1289.5; DB 1; Length 334;
Best Local Similarity 74.3%; Pred. No. 3.4e-82;
Matches 249; Conservative 33; Mismatches 52; Indels 1; Gaps 1;
CC 2 VVKVINGFGRIGRLAFRRIONVEGEVETRLNDLTDPNMLAHLKDYDTTQGRFDGTVEV 61
CC 1 MTKVAINGFGRIGRLALRRILEVPGLEVVAINDLTDKMLAHLKDYDSSQGRFNGEIEVK 60
CC 62 EGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 121
CC 61 EGAFVNGKEVKVFAEADPEKLPWGELGIDVLECTGFFTKKEAEHVRA-GAKKVIS 119
CC 122 APGNDVKTIVFNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIIA 181
CC 120 APAGNDLKTIVFNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIEKGFMTTIIA 179
CC 182 TGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTG 241
CC 180 TNDQNTLDGPHRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTG 239
CC 242 SVTELVTLTKNVSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 301
CC 240 SVTELVTLTKNVSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 299
CC 302 VDGSQLVKVSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 336
CC 300 VDGSQLVKVSVDENAMKAASNDSEFGYTEDPIVSSDIVGVSGLFDATQTKVM 334
CC -----
RESULT 6
G3P_CLOAB STANDARD; PRT; 334 AA.
ID G3P_CLOAB
```


Best Local Similarity 69.4%; Pred. No. 2.6e-75;
Matches 234; Conservative 39; Mismatches 62; Indels 2; Gaps 2;
QY 1 MVVKVINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKDYDTQGRFDGTVEV 60
Db 1 MAIKVAINGFGRIGRLAFRIQNVGVEVAVNDLTPNMLAHLKDYDTQGRFTGEVEV 60
QY 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 IEGFRVNGKEIKSFDEPDAGKLPWGLDLDIVVLECTGYTDKEKAQAHIDA-GAKKVI 119
QY 121 TAPGNDVKTIVFNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 120 SAPAKGDVKTIVFNTNHDLDGSETVVGASCTTNLAPVAKVLSDFGLVEGFMTTIHA 179
QY 181 YTGDMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 180 YTGDMQNTQDAPHRKGDKRRARAAENIIPNSTGAAGAIGLVIPIDGKLDGAAQRPVPT 239
QY 241 GSVTELVVTLTK-NVSVDEINAAKMAASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKV 299
Db 240 GSVTELVVTLTKQDVTVQVNEAMKNASNESFYTEDEIVSSDIVGVSGLFDTATQTRV 299
QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 300 MTVGDRQLVKVAAWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 8

G3P1_STAAM STANDARD; PRT; 336 AA.
AC Q925C5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH 1).
GN GAPA OR GAP OR SAV0772 OR SA0727 OR MW0734.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315),
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BB;
RA Morrissey J.A., Williams P.;
RT "Isolation and characterisation of a glycolytic operon in
RT Staphylococcus aureus";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";

RL Lancet 359:1819-1827 (2002).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ133520; CAB38645.1; -.
DR EMBL; AP003360; BAB56934.1; -.
DR EMBL; AP003131; BAB41960.1; -.
DR EMBL; AP004824; BAB94599.1; -.
DR PIR; E89850; E89850.
DR SWISS-2DPAGE; Q925C5; STAA.
DR HSSP; P17721; 1HDG.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT BINDING 151 151
FT ACT_SITE 178 178
FT ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
SQ SEQUENCE 336 AA; 36281 MW; 37A6CEA9376779E5 CRC64;

Query Match 69.0%; Score 1183; DB 1; Length 336;
Best Local Similarity 68.2%; Pred. No. 7.9e-75;
Matches 230; Conservative 43; Mismatches 62; Indels 2; Gaps 2;
QY 1 MVVKVINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKDYDTQGRFDGTVEV 60
Db 1 MAVKVAINGFGRIGRLAFRIQNVGVEVAVNDLTPNMLAHLKDYDTQGRFTGEVEV 60
QY 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 VDGFRVNGKEIKSFSEPDASKLPWKDLNIDVLECTGYTDKKAQAHIEA-GAKKVI 119
QY 121 TAPGNDVKTIVFNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 120 SAPATGDLKTIVFNTNHDLDGSETVVGASCTTNLAPVAKVLDNDGFLVEGLMTTIHA 179
QY 181 YTGDMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 180 YTGDMQNTQDAPHRKGDKRRARAAENIIPNSTGAAGAIGLVIPIDGKLDGAAQRPVPT 239
QY 241 GSVTELVVTLTK-NVSVDEINAAKMAASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKV 299
Db 240 GSVTELVVTLTKQDVTVQVNEAMKNASNESFYTEDEIVSSDIVGVSGLFDTATQTRV 299
QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 300 MSVGDRLVKVAAWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 9

G3P1_LACDE STANDARD; PRT; 338 AA.
ID ID_G3P_LACDE
AC O32755;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP.
OS Lactobacillus delbrueckii (subsp. bulgaricus).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1585;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B107;
RX MEDLINE=98240227; PubMed=9579064;
RA Branny P., Delatorre F., Garel J.R.;
RT "An operon encoding three glycolytic enzymes in Lactobacillus
RT delbrueckii subsp. bulgaricus: glyceraldehyde-3-phosphate
RT dehydrogenase, phosphoglycerate kinase and triosephosphate
RT isomerase.";
RL Microbiology 144:905-914(1998).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ000339; CAA04014.1; --
CC DR PIR; T09633; T09633.
CC DR HSSP; P17721; 1HDG.
CC DR InterPro; IPR000173; GAP dhdrogenase.
CC DR InterPro; IPR006424; GAPDH-I.
CC DR Pfam; PF00044; gpdh; 1.
CC DR Pfam; PF02800; gpdh C; 1.
CC DR PRINTS; PR00078; G3PDHDRGNASE.
CC DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC DR PROSITE; PS00071; GAPDH; FALSE_NEG.
KW Glycolysis; NAD; Oxidoreductase.
FT BINDING 156 156 GLYCERALDEHYDE 3-PHOSPHATE
FT (BY SIMILARITY).
FT ACT_SITE 183 183 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
FT SEQUENCE 338 AA; 36564 MW; DF97E8C9CA4F7DFA CRC64;
Query Match 58.5%; Score 1003; DB 1; Length 338;
Best Local Similarity 58.7%; Pred. No. 2.2e-62;
Matches 199; Conservative 45; Mismatches 89; Indels 6; Gaps 3;
QY 1 MVVKVINGFGRIQLAFRI---QNVGVEVTRINDLTDPNMLAHLKLYDTTQGRFDG 56
DB 1 MTVKIGINGFGRIQLAFRRIMDLGEETKDIEVVAINDLTPAMLAHLKLYDSTHGTFDH 60
QY 57 TVEVKEGFEVNGNFIKVSARDPENIDWA-TDGVIEIVLEATGFFAKKEAEKHLHANGA 115
DB 61 EVSATEDSLVVDGKKYRVYAEPAQNIPIWVKNVDGVDFVLECTGFYTSKAKSQAHLDA-GA 119
QY 116 KKVITAPGNDYKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDFAFGIQGLM 175
DB 120 KRVLIAPAGNDLKTIVSVNQDTLTADTIVSAGSCTTNLAPMANALNKEFGIQVGTM 179
QY 176 TTIHAYTGDMILGPHRGGLRRRAGAAINVPNSTGAAKAIGLVIPELNGKLDGAQR 235
DB 180 TTIHAYTATQKVLGDPDRGNFRNRAAENIIPHSTGAAKAIGLVLPELNGKLDGHAQR 239
QY 236 VPVPTGVTVELVTLDKNVSVDEINAAKASNDISFGYTEDPIVSSDIVGSYGLFDAT 295
DB 240 VPVKDGSSETELVTLDKKVTAEVNAAMKKYSPSFAYNADQIVSTDVLGMTAGSIFDPT 299

QY 296 QTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTEYFAKI 334
DB 300 QTQVITAGDKQLVKTVAWYDNEYSFCTCMVRLHLHFATL 338
RESULT 10
G3P3 ECO57
ID -G3P3 ECO57 STANDARD; PRT; 333 AA.
AC P58072;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).
GN GAPC OR Z2304 OR ECS2022.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimailanta E.T., Potamouisis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB005364; AAG56359.1; --
CC EMBL; AP002557; BAB35445.1; --
CC PIR; C85737; C85737.
CC PIR; F90881; F90881.
CC HSSP; P17721; 1HDG.
CC InterPro; IPR000173; GAP dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE
FT (BY SIMILARITY).

	FT	ACT_SITE	177	177	ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).					
	FT	SEQUENCE	333 AA;	35763 MW;	A2F77CB2E773E64C CRC64;					
	SQ	Query Match	57.9%;	Score 993;	DB 1; Length 333;					
		Best Local Similarity	59.3%;	Pred. No.	1.le-61;					
		Matches 198;	Conservative	54;	Mismatches 76;	Indels	6;	Gaps	5;	
QY	4	KVINGPGRIGRLAFRRIONVE-GVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEVKE	122							
Dd	3	KVGINGPGRIGRLVLRLLEVKSNIDWAINDLTSRKILAYLLKHDSNYGPFWSVDLTE	62							
QY	63	GGFEVNGNFIKVSAERDPEINIDWATDGVEIIVLEATGFPAKKEAAEKHLHANGAKKWITA	122							
Dd	63	DSLIVDKSIAVYAEEKAENIPWKAGAEIIIVECTGFTYSAEKSQAHLDA-GAKKVLISA	121							
QY	123	PGGNVDKVTVFNTNHDI LDGTETVISGASCTTNCLAPMAKALHD AFGIQGLMTTI HAYT	182							
Dd	122	PAG-EMKIIVYKVNDDTL DGNNDIVSVASCTTNCLAPYAKALHDSFGIEVGVTTIHAYT	180							
QY	183	GDQMILDGPHRGDDLRRRAGAANIVPNSTGA KA IGLVIP ELNGKL DGAAQRVPVPTGS	242							
Dd	181	GTQSLVDGP-RGXDLRASRAAENIIPH TTGA KA IGLVIP ELSGKLKGHAORVPVK TGS	239							
QY	243	VTELVVTLDKNVSVD EINAMKAA--SNSDFGYTEDPI VSSDIVGVSYGSLF DATQT KVM	300							
Dd	240	VTELVSILCKKVTAEEVNNA LKQATTNNESFGYT DEEVSSDI IGSHFGSV FDATQ TEIT	299							
QY	301	EVDGSQLVKVSWYDNEMS YTAQLVRTLEYFA KI	334							
Dd	300	AVGDLQLVKTVAWYDNEY GFVTLRTLEKF AKL	333							

RESULT 11

```

G3P3 ECOLI STANDARD; PRT: 333 AA.
AC P33898; P76094; P78062; P78291; Q03850; Q63208;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).
GN GAPC OR B1416/B1417.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=River isolate, and Clinical isolate;
RX MEDLINE=98283700; PubMed=96223357;
RA Espinosa-Urgel M., Kolter R.;
RT "Escherichia coli genes expressed preferentially in an aquatic environment.";
RL Mol. Microbiol. 28:325-332(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Hidalgo E., Limon A., Aguilar J.;
RT "A second Escherichia coli gene with similarity to gapA." ;
RL Microbiology 12:99-106(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12." ;
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
    
```


DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHNRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE
FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS
FT CONFLICT 39 39 K -> Y (IN REF. 7 AND 8).
FT SEQUENCE 333 AA; 35649 MW; E88223297376B0A0 CRC64;
SQ
Query Match 57.8%; Score 992; DB 1; Length 333;
Best Local Similarity 59.3%; Pred. No. 1.2e-61;
Matches 198; Conservative 54; Mismatches 76; Indels 6; Gaps 5;
QY 4 KVGINGFGRIGRLAFRRIONVE-GVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEVKE 62
DB 3 KVGINGFGRIGRLVGLRLEKSNIDVVAINDLTSPKILAYLLKHDSDNYGPPFWSVDFT 62
QY 63 GGFVNGNFIVKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVVITA 122
DB 63 DSLIVDGKSIAYAEKEAKNIPWAKAGAEIIVCTGFTYSAEKSQAHLDA-GAKKVLISA 121
QY 123 PGGNDVKTVVNTNHDILDGTETVISGACTTNCCLAPMAKALHDAFGIOKGLMTTIHAYT 182
DB 122 PAG-EMKTIIVNVNDTLDGNDTIIVSVASCTTNCCLAPMAKALHDSFGIEVGTMTTIHAYT 180
QY 183 GDQMILDPHRRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGS 242
DB 181 GTQSLVDGP-RGDLRASRAAAENIIPHTTGAAGAIGLVIPELNGKLDGAAQRPVPTGS 239
QY 243 VTELVTLDKNVSVDEINAAKAA--SNDSPGYTEDPIVSSDIVGSYGSGLFDATQTKVM 300
DB 240 VTELVSILGKVTABEVNNAKQATTNNESTGYTDEEIVSSDIIGSHFGSVFDTQTEIT 299
QY 301 EVDGSQLKVVSWYDNEMSYTAQLVRLTLEYPAKI 334
DB 300 AVGDQLVKTVMYDNEYGVFTQLIRLTLEKFAKL 333

RESULT 12
G3P_MYCPN STANDARD; PRT; 337 AA.
AC P75358;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR GAP OR MPN430 OR MP411.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelteich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE000040; AAB96059.1; -.
DR PIR; S73737; S73737.
DR HSSP; P17721; 1HDG.
DR InterPro; IPR000173; GAP dhhydrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHNRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE
FT ACT_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS
FT SEQUENCE 337 AA; 36805 MW; 550747A529ABCA83 CRC64;
SQ

Query Match 53.7%; Score 921; DB 1; Length 337;
Best Local Similarity 54.1%; Pred. No. 1e-56;
Matches 180; Conservative 54; Mismatches 95; Indels 4; Gaps 3;
QY 3 KVGINGFGRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEVKE 62
DB 8 IRVAINGFGRIGRLVFRALLSQKNIIEIVAVNDLTHPDTLAHLKDYDSAHEGFFKKVAKD 67
QY 63 GGFVNGNFIVKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVVITA 122
DB 68 NTLMDVKKVLFSEKDPANLPWAENHNDIVVESTGRFVSEEGASLHQA-GAKRVIISA 126
QY 123 PG-GNDVKTVVNTNHDILDGTETVISGACTTNCCLAPMAKALHDAFGIOKGLMTTIHAY 181
DB 127 PAKQKTIKTVVYVNVNKKIINAEDKIIISAASCTNCLAPMVHVKFNFGILHGTMTVTVHAY 186
QY 182 TGDQMILDPHRRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTG 241
DB 187 TADQRLQDAPH--SDLRRARAAACNIVPTTGAAGAIGLVPVPEATGKNGMALRVPVLTG 244
QY 242 SVTELVTLDKNVSVDEINAAKAAASNDSPGYTEDPIVSSDIVGSYGSGLFDATQTKVME 301
DB 245 SIVELCVALEKDATVEQINQAMKKAASAFRYCEDEIVSSDIVGSEHGSIFDSKLTNIIIE 304
QY 302 VDGSQLKVVSWYDNEMSYTAQLVRLTLEYPAKI 334
DB 305 VDGKLYKYAWYDNESYVNVQLVRVNVYCAKL 337

RESULT 13
G3P_CORGL STANDARD; PRT; 334 AA.
AC Q01551;
DT 01-APR-1993 (Rel. 25, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR CGL1588.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13059 / AS019;
RX MEDLINE=93015645; PubMed=1400158;
RA Eikmanns B.J.;
RT "Identification, sequence analysis, and expression of a
Corynebacterium glutamicum gene cluster encoding the three glycolytic
enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate
kinase, and triosephosphate isomerase";

RL J. Bacteriol. 174:6076-6086(1992).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;

RA Nakagawa S.;

RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -|- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +

CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -|- PATHWAY: Second phase of glycolysis; first step.

CC -|- SUBUNIT: Homotetramer.

CC -|- SUBCELLULAR LOCATION: Cytoplasmic.

CC -|- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate

CC dehydrogenase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X59403; CAA42045.1; -.

CC EMBL; AP005279; BAB98981.1; -.

CC PIR; A43260; A43260.

CC HSSP; P00362; 1GD1.

CC InterPro; IPR000173; GAP_dhhydrogenase.

CC InterPro; IPR006424; GAPDH-I.

CC Pfam; PF00044; gpdh; 1.

CC PRINTS; PR00078; G3PDHDRGNASE.

CC TIGRFAMS; TIGR01534; GAPDH-I; 1.

CC PROSITE; PS00071; GAPDH; 1.

CC Glycolysis; Oxidoreductase; NAD; Complete proteome.

FT BINDING 153 153 GLYCERALDEHYDE 3-PHOSPHATE.

FT ACT SITE 180 180 ACTIVATES THIOL GROUP DURING CATALYSIS.

FT CONFLICT 25 26 SD -> NG (IN REF. 1).

FT CONFLICT 333 334 KL -> QALN (IN REF. 1).

SQ SEQUENCE 334 AA; 36045 MW; 33792AF65FA90FA7 CRC64;

Query Match 53.0%; Score 908.5; DB 1; Length 334;

Best Local Similarity 56.2%; Pred. No. 7.4e-56;

Matches 190; Conservative 41; Mismatches 94; Indels 13; Gaps 6;

QY 1 MVKVGINGFGRIGRLAFRR1-QNVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVE 59

Db 1 MTIRVGINGFGRIGRNFRAVLESDDEVVAVNDLTDNKTSLTLKFPDSIMGRLGQVE 60

QY 60 VKGGFEVNGNFIKVSABERDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKVV 119

Db 61 YDDDSITVGGKRIAVYAERDPKNDWAHNVDIVESTGFTDANAARAKAHIEA-GAKKVI 119

QY 120 ITAPGNDVKTVVNTNHDILD-GTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTI 178

Db 120 ISAPASNEATFVYGVNHESYDPENHNVISGASCTTNCLAPMAKVLNDRKFGIENGLMTTV 179

QY 179 HAYTGDQMILDGPHRGGLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPV 238

Db 180 HAYTGDQRLHDAPHR--DLRRARAAAVNIPTSTGAAKAVALVLPKLGKLDGYALRPV 237

QY 239 PTGSVTELVTLDKNVSVDEINAAMKAASNDSFG---YTEDPIVSSDIYGVSYGSLFDA 294

Db 238 ITGSATDLTFNTKSEVTVESINAAIKEAAVGEFGETLAYSEELVSTDIVHDSHGSI 297

QY 295 TQTKMEVDGSQLVKVSVYDNEMSYTAQLVRLTEYFA 332

Db 298 GLTKV----SGNTVKVSVYDNEMGYTCQLRLRLTELVA 331

RESULT 14

G3P_MYCGE

ID G3P_MYCGE STANDARD; PRT; 337 AA.

AC P47543;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)

DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).

GN GAPA OR GAP OR MG301.

OS Mycoplasma genitalium.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2097;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33530 / G-37;

RX MEDLINE=96026346; PubMed=7569993;

RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,

RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,

RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,

RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,

RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,

RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;

RT "The minimal gene complement of Mycoplasma genitalium.";

RL Science 270:397-403(1995).

RN [2]

RP SEQUENCE OF 1-81 AND 279-337 FROM N.A.

RC STRAIN=ATCC 33530 / G-37;

RX MEDLINE=94075230; PubMed=8253680;

RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;

RT "A survey of the Mycoplasma genitalium genome by using random

RT sequencing.";

RL J. Bacteriol. 175:7918-7930(1993).

CC -|- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +

CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -|- PATHWAY: Second phase of glycolysis; first step.

CC -|- SUBUNIT: Homotetramer (By similarity).

CC -|- SUBCELLULAR LOCATION: Cytoplasmic.

CC -|- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate

CC dehydrogenase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U39710; AAC71523.1; -.

CC EMBL; U02213; AAD12507.1; -.

CC EMBL; U02178; AAD12463.1; -.

CC PIR; C64233; C64233.

CC HSSP; P17721; 1HDG.

CC TIGR; MG301; -.

CC InterPro; IPR000173; GAP_dhhydrogenase.

CC InterPro; IPR006424; GAPDH-I.

CC Pfam; PF00044; gpdh; 1.

CC PRINTS; PR00078; G3PDHDRGNASE.

CC TIGRFAMS; TIGR01534; GAPDH-I; 1.

CC PROSITE; PS00071; GAPDH; 1.

CC Glycolysis; Oxidoreductase; NAD; Complete proteome.

FT BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE

FT ACT_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS

FT (BY SIMILARITY).

FT (BY SIMILARITY).

SQ SEQUENCE 337 AA; 37097 MW; FALEA1966687006B CRC64;

Query Match 52.9%; Score 908; DB 1; Length 337;

Best Local Similarity 53.2%; Pred. No. 8.2e-56;

Matches 177; Conservative 54; Mismatches 98; Indels 4; Gaps 3;

QY 3 VKVGINGFGRIGRLAFRR1QNVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEVKE 62

Db 8 IKVAINGFGRIGRLVFRSLSKANVEVVAINDLTQPEVLAHLKYDSANGELKRRKITVKQ 67

QY 63 GGFEVNGNFKVSAERDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVITA 122
Db 68 NILQIDRKVVVFSEKDPQNLDPWDEHDIDVVIESTGRFVSEGAHLKA-GAKRVIISA 126
QY 123 PG-GNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTHAY 181
Db 127 PAKEKTIRTVVYNNVNHKTISSDDKIISAASCTTNCLAPLVHLEKNFGIVYGTMLTVHAY 186
QY 182 TGDQMILDGPHRGGLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTG 241
Db 187 TADQRLQDAPH--NDLRRARAAAANIVPTTTGAAKAIGLVPEANGKLGMSLRVPLTG 244
QY 242 SVTELVTLDKNVSVDEINAAMKAASNDISFGYTEDPIVSSDIVGVSGLSFDATQTKVME 301
Db 245 SIVELSVVLEKSPSVEQVNQAMKRFASAFKYCEDPIVSSDVVSSEYGSIFDSKLTNIVE 304
QY 302 VDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKI 334
Db 305 VDKMKLYKVAYWYDNESYVHQLVVRVVSYCAKL 337
RESULT 15
G3P1_BACSU STANDARD; PRT; 334 AA.
AC P09124;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glycerinaldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH) (NAD-
DE dependent glycerinaldehyde-3-phosphate dehydrogenase).
GN GAPD OR GAP OR BSU33940.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BDI70;
RX MEDLINE=89160255; PubMed=2493629;
RA Viaene A., Dhaese P.;
RT "Sequence of the glycerinaldehyde-3-phosphate dehydrogenase gene from
RT Bacillus subtilis."
RL Nucleic Acids Res. 17:1251-1251(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.P., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Frit C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE OF 1-30.
RC STRAIN=168 / JH642;
RX MEDLINE=96345629; PubMed=8755892;
RA Graumann P., Schroeder K., Schmid R., Marahiel M.A.;
RT "Cold shock stress-induced proteins in Bacillus subtilis.";
RL J. Bacteriol. 178:4611-4619(1996).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20261518; PubMed=10799476;
RA Fillingner S., Boschi-Muller S., Azza S., Dervyn E., Branlant G.,
RA Aymerich S.;
RT "Two glycerinaldehyde-3-phosphate dehydrogenases with opposite
RT physiological roles in a nonphotosynthetic bacterium.";
RL J. Biol. Chem. 275:14031-14037(2000).
CC -!- FUNCTION: More active in catabolism.
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glycerinaldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13011; CAA31434.1; --
DR EMBL; Z99121; CAB15399.1; --
DR PIR; S02754; DEBSG.
DR HSSP; P00362; 1GD1.
DR Subtilist; BG10827; gapA.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
FT INIT MET 0
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 334 AA; 35701 MW; 1283D3E6CF5095EC CRC64;
Query Match 52.0%; Score 891.5; DB 1; Length 334;
Best Local Similarity 54.5%; Pred. No. 1.1e-54;
Matches 183; Conservative 41; Mismatches 101; Indels 11; Gaps 5;
QY 3 VKVGINGFGRIGRLAFRRIONVEGVETRIINDLTDPNMLAHLKVDYDTQGRFDGTVEVKE 62
Db 2 VKVGINGFGRIGRNVFRAALNNPEVEVAVNDLTDANMLAHLLOYDSVHGKLDAEVSVDG 61
QY 63 GGFEVNGNFKVSAERDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVITA 122
Db 62 NNLVNGKTIIEVSAERDPAKLSWGKQGVVEIVVESTGFTTKRADAKHLEA-GAKKVIISA 120
QY 123 PGNDVKTVVFNTHDILD-GTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTHAY 181
Db 121 PANEEDITIVMGVNEKDYDAANHDVSNASCTTNCLAPFAKVLNDKFGIKRGMVTVHSY 180
QY 182 TGDQMILDGPHRGGLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTG 241
Db 181 TNDQQLDLPHK--DYRRARAAAENIPTSTGAAKAVSLVLPKLGKLGNGGMRVPTPNV 238

Search completed: May 11, 2004, 16:48:17
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: May 11, 2004, 16:45:16 ; Search time 46 Seconds
(without alignments)
2304.655 Million cell updates/sec

Title: US-10-650-369-12
Perfect score: 1715
Sequence: 1 MVVKVGINFGFRIGRLAFRR.....EMSYTAQLVRLTLEYFAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1715	100.0	336	Q83ZF7	Q83zf7 streptococc
2	1618	94.3	336	Q8KVU6	Q8kvu6 streptococc
3	1609	93.8	336	Q9ALW2	Q9alw2 streptococc
4	1609	93.8	336	Q8E3E8	Q8e3e8 streptococc
5	1609	93.8	336	Q8DXS8	Q8dxs8 streptococc
6	1596	93.1	336	Q8GCR7	Q8gcr7 streptococc
7	1593.5	92.9	335	Q97NL1	Q97nl1 streptococc
8	1593.5	92.9	359	Q8CWN6	Q8cwn6 streptococc
9	1584	92.4	336	Q8KHG1	Q8khg1 streptococc
10	1579.5	92.1	337	Q8DVV3	Q8dvv3 streptococc
11	1578	92.0	336	Q8KVU5	Q8kvu5 streptococc
12	1572	91.7	336	Q8VVB9	Q8vvb9 streptococc
13	1518	88.5	320	Q9LSX6	Q9lsx6 streptococc
14	1468.5	85.6	309	Q9AJT7	Q9ajt7 streptococc
15	1467.5	85.6	309	Q9AUT9	Q9aut9 streptococc
16	1465.5	85.5	309	Q9AJT4	Q9ajt4 streptococc

17	1463.5	85.3	309	2	Q9AJT5	Q9ajt5 streptococc
18	1462.5	85.3	308	2	Q9AJT8	Q9ajt8 streptococc
19	1445.5	84.3	305	2	Q9AJT6	Q9ajt6 streptococc
20	1410	82.2	336	16	Q9CDH4	Q9cdh4 lactococcu
21	1337.5	78.0	333	16	Q833I8	Q833i8 enterococcu
22	1262	73.6	336	16	Q8Y4I1	Q8y4i1 listeria m
23	1261	73.5	336	16	Q928H9	Q928h9 listeria in
24	1227.5	71.6	332	16	Q8XKT9	Q8xkt9 clostridium
25	1218.5	71.0	335	16	Q8RFN9	Q8rfn9 fusobacteri
26	1197.5	69.8	334	16	Q9JW8	Q9jw8 neisseria m
27	1190	69.4	336	16	Q8CPY5	Q8cp5 staphylococ
28	1189.5	69.4	330	2	Q83UX3	Q83ux3 neisseria g
29	1186.5	69.2	330	2	Q84HZ6	Q84hz6 neisseria g
30	1185.5	69.1	330	2	Q84HZ4	Q84hz4 neisseria g
31	1184.5	69.1	334	16	Q9JX95	Q9jx95 neisseria m
32	1183	69.0	336	16	Q9Z5C5	Q9z5c5 staphylococ
33	1182.5	69.0	330	2	Q84HZ5	Q84hz5 neisseria g
34	1178.5	68.7	335	2	Q93M61	Q93m61 staphylococ
35	1138	66.4	311	2	Q8L2R0	Q8l2r0 staphylococ
36	1132.5	66.0	310	2	Q8L2P7	Q8l2p7 staphylococ
37	1124.5	65.6	310	2	Q8L2P6	Q8l2p6 staphylococ
38	1124	65.5	311	2	Q8L2P8	Q8l2p8 staphylococ
39	1123.5	65.5	310	2	Q8L2Q6	Q8l2q6 staphylococ
40	1121.5	65.4	311	2	Q8VM79	Q8vm79 rhodospiril
41	1120.5	65.3	310	2	Q8L2Q3	Q8l2q3 staphylococ
42	1120.5	65.3	310	2	Q8L2P5	Q8l2p5 staphylococ
43	1118.5	65.2	310	2	Q8L2Q2	Q8l2q2 staphylococ
44	1116.5	65.1	310	2	Q8L2R7	Q8l2r7 staphylococ
45	1115.5	65.0	310	2	Q8L2R2	Q8l2r2 staphylococ

ALIGNMENTS

RESULT 1

Q83ZF7 PRELIMINARY; PRT; 336 AA.
ID Q83ZF7
AC Q83ZF7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Glyceraldehyde 3-P dehydrogenase.
GN GAPC.
OS Streptococcus dysgalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1334;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Casal J.F., Bolton A., Song X.-M., Willson P., Potter A.A.;
RT "Use of the surface proteins GapC and Mig of Streptococcus
RT dysgalactiae as protective antigens against mastitis in non-lactating
RT cows";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF375662; AAP31408.1;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
SQ SEQUENCE 336 AA; 35928 MW; 11828218CF037076 CRC64;

Query Match 100.0%; Score 1715; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 5.7e-107;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MVVKVGINFGFRIGRLAFRRIONVEGVETRLNDLTPNMLAHLLKYDTTQGRFDGTVEV	60
Db	1	MVVKVGINFGFRIGRLAFRRIONVEGVETRLNDLTPNMLAHLLKYDTTQGRFDGTVEV	60

QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
DB 121 TAPGNDVKTIVFNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDENIINAAKKAASNDSPGYTDPVSSDIVGVSGLFDATQTKVM 300
DB 241 GSVTELVVTLTKNVSVDENIINAAKKAASNDSPGYTDPVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 2

Q8KVU6 ID Q8KVU6 PRELIMINARY; PRT; 336 AA.
AC Q8KVU6;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
OS Streptococcus uberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1349;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 9927;
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421900; AAM73771.1; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHRCNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 35906 MW; 755D74A4548E61D4 CRC64;

Query Match 94.3%; Score 1618; DB 2; Length 336;
Best Local Similarity 92.9%; Pred. No. 1.8e-100;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180

DB 121 TAPGDDVKTIVFNTNHDILDTGTTVSGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDENIINAAKKAASNDSPGYTDPVSSDIVGVSGLFDATQTKVM 300
DB 241 GSVTELVVTLTKNVSVDENIINAAKKAASNDSPGYTDPVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
DB 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 3

Q9ALW2 ID Q9ALW2 PRELIMINARY; PRT; 336 AA.
AC Q9ALW2;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J48;
RA Seifert K.N., Bleiweis A.S., McArthur W.P., Brady L.J.;
RT "The Group B Streptococcal Surface Antigen Delta is a Glyceraldehyde
RT 3-Phosphate Dehydrogenase."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF338416; AAK14387.1; -.
DR HSSP; P00362; IGD1.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHRCNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 93.8%; Score 1609; DB 2; Length 336;
Best Local Similarity 92.0%; Pred. No. 7.2e-100;
Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
DB 121 TAPGNDVKTIVFNTNHDILDTGTTVSGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDENIINAAKKAASNDSPGYTDPVSSDIVGVSGLFDATQTKVM 300

Db 241 GSVTELVATLEKDVTVVEEVNAAMKAAANDSYGYTEDPIVSSDIVGISYGSFLFDATQTKVQ 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIAK 336

RESULT 4

Q8E3E8 PRELIMINARY; PRT; 336 AA.
AC Q8E3E8;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN GBS1811.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766853; CAD47470.1; --
DR Sagalists; gbs1811; --
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh.C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 93.8%; Score 1609; DB 16; Length 336;
Best Local Similarity 92.0%; Pred. No. 7.2e-100;
Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

QY 1 MVVKVINGFGFRIGRLAFRRIONVEGVETVRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGFRIGRLAFRRIONVEGVETVRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFTKVSARDPENIDWATDGVVEIVLEATGFFAKAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEATGFFASKEAEQHIHENGAKKVI 120
QY 121 TAPGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAANIVNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLTKNVSVDENAAKAAANDSGYTEDPIVSSDIVGSYGSFLFDATQTKVM 300
Db 241 GSVTELVATLEKDVTVVEEVNAAMKAAANDSYGYTEDPIVSSDIVGSYGSFLFDATQTKVQ 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIAK 336

RESULT 6

RESULT 5
Q8DXS8 PRELIMINARY; PRT; 336 AA.
AC Q8DXS8;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN GAP OR SAG1768.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae."
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AE014272; AAN00631.1; --
DR TIGR; SAG1768; --
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh.C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 93.8%; Score 1609; DB 16; Length 336;
Best Local Similarity 92.0%; Pred. No. 7.2e-100;
Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

QY 1 MVVKVINGFGFRIGRLAFRRIONVEGVETVRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGFRIGRLAFRRIONVEGVETVRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFTKVSARDPENIDWATDGVVEIVLEATGFFAKAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEATGFFASKEAEQHIHENGAKKVI 120
QY 121 TAPGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAANIVNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLTKNVSVDENAAKAAANDSGYTEDPIVSSDIVGSYGSFLFDATQTKVM 300
Db 241 GSVTELVATLEKDVTVVEEVNAAMKAAANDSYGYTEDPIVSSDIVGSYGSFLFDATQTKVQ 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIAK 336

```
Q8GCR7
ID Q8GCR7 PRELIMINARY; PRT; 336 AA.
AC Q8GCR7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).
OS Streptococcus suis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1307;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S735;
RA Brassard J., Gottschalk M., Quesy S.;
RT "Cloning and purification of Streptococcus suis serotype 2
glyceraldehyde-3-phosphate dehydrogenase.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY167026; AAN86058.1; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000173; GAP_dhrogenase.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF02800; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 35825 MW; 1D8B9B1A492DCF59 CRC64;

Query Match 93.1%; Score 1596; DB 2; Length 336;
Best Local Similarity 91.1%; Pred. NO. 5.4e-99;
Matches 306; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIQLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIQLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120

QY 121 TAPGNDVKTVEFNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 121 TAPGNDVKTVEFNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHG 180

QY 181 YTGDMILDGPHRGGLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGGLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

QY 241 GSVTELVTLTKNSVDEINAAKASNDSTFGYTEDPIVSSDIVGSGSLFDATQTKVM 300
Db 241 GSVTELVTLTKNSVDEINAAKASNDSTFGYTEDPIVSSDIVGSGSLFDATQTKVI 300

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 7
Q97NL1 PRELIMINARY; PRT; 335 AA.
ID Q97NL1;
AC Q97NL1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12).
GN SP2012 OR GAPDH.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
```

```
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D., C.L.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
pneumoniae.";
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC11733;
RA Bergmann S., Hammerschmidt S.;
RT "Identification of pneumococcal GAPDH as plasmin(ogen)-binding
protein.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007490; AAK76079.1; -.
DR EMBL; AJ505822; CAD44376.1; -.
DR PIR; F95235; F95235.
DR HSSP; P00354; 3GPD.
DR TIGR; SP2012; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 335 AA; 35856 MW; DA483CEA423E747B CRC64;

Query Match 92.9%; Score 1593.5; DB 16; Length 335;
Best Local Similarity 92.6%; Pred. NO. 7.9e-99;
Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 1 MVVKVINGFGRIQLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIQLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120

QY 121 TAPGNDVKTVEFNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 121 TAPGNDVKTVEFNTNHDILDTGTETVISGASCTTNCLAPMAKALQDNFVVEGLMTTIHA 179

QY 181 YTGDMILDGPHRGGLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 180 YTGDMILDGPHRGGLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGSAQRPVPT 239

QY 241 GSVTELVTLTKNSVDEINAAKASNDSTFGYTEDPIVSSDIVGSGSLFDATQTKVM 300
Db 240 GSVTELVAVLEKNVTDEVNAAKASNESYGTEDPIVSSDIVGMSYGSLEFDATQTKVL 299

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 300 DVDGKQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 8
Q8CWN6 PRELIMINARY; PRT; 359 AA.
ID Q8CWN6
AC Q8CWN6;
```

DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (Phosphorylating)
DE (EC 1.2.1.12).
GN GAPA OR SPR1825.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
McAhen S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008547; AAL00628.1; -;
DR PIR; G98099; G98099.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 359 AA; 38763 MW; 61EF0E375B330B60 CRC64;

Query Match 92.9%; Score 1593.5; DB 16; Length 359;
Best Local Similarity 92.6%; Pred. No. 8.7e-99;
Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;
QY 1 MVVKVINGFGFRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 25 MVVKVINGFGFRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 84
QY 61 KEGGFEVNGNFKVSAERDPENIDWATDGVETVLEATGFFAKKAEKHLHANGAKKVI 120
Db 85 KEGGFEVNGKFKVSAERDPQIDWATDGVETVLEATGFFAKKAEKHL-KGGAKKVI 143
QY 121 TAPGGNDVKTVVNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 144 TAPGGNDVKTVVNTNHDLDGTETVISGASCTTNCLAPMAKALQDNFGVWEGMLTTIHA 203
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 204 YTGDMILDGPHRGDLRRARAGAAANIVNSTGAAGAIGLVIPELNGKLDGSAQRPVPT 263
QY 241 GSVTELVVTLTKNVSVDEINAAKAAASNDSPGYTEDPIVSSDIVGSYSLFDATQTKVM 300
Db 264 GSVTELVAVLEKNVTVDENVAAKAAASNESYGYTEDPIVSSDIVGSYSLFDATQTKVL 323
QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTEYFAKIAK 336
Db 324 DVDGKQLVKVSVWYDNEMSYTAQLVRLTEYFAKIAK 359

RESULT 9
Q8KHG1
ID Q8KHG1 PRELIMINARY; PRT; 336 AA.
AC Q8KHG1;
DT 01-OCT-2002 (TReMBLrel. 22, Created)

DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus agalactiae, and
OS Streptococcus iniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311, 1346;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.agalactiae, and S.iniae; STRAIN=ATCC 27541, and 9117;
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
antigens confers protection against heterologous challenge with
Streptococcus uberis";
RT Streptococcus uberis";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
DEHYDROGENASE FAMILY.
DR EMBL; AF421899; AAM73770.1; -;
DR EMBL; AF421902; AAM73773.1; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 35723 MW; ABAB9E14F3ED1111 CRC64;

Query Match 92.4%; Score 1584; DB 2; Length 336;
Best Local Similarity 90.5%; Pred. No. 3.4e-98;
Matches 304; Conservative 19; Mismatches 13; Indels 0; Gaps 0;
QY 1 MVVKVINGFGFRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGFRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFKVSAERDPENIDWATDGVETVLEATGFFAKKAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGSFVKVSAEREPANIDWATDGVETVLEATGFFAKKAAAEQHIHANGAKKVI 120
QY 121 TAPGGNDVKTVVNTNHDLDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGNDVKTVVNTNHDLDGTETVISGASCTTNCLAPMAKALQDNFGVWEGMLTTIHA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAANIVNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDEINAAKAAASNDSPGYTEDPIVSSDIVGSYSLFDATQTKVM 300
Db 241 GSVTELVAVLEKDTSVVEINAAKAAANDSYGYTEDAIVSSDIVGSYSLFDATQTKVQ 300
QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTEYFAKIAK 336
Db 301 TVDGNQLVKVSVWYDNEMSYTAQLVRLTEYFAKIAK 336

RESULT 10
Q8DVV3
ID Q8DVV3 PRELIMINARY; PRT; 337 AA.
AC Q8DVV3;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Extracellular glyceraldehyde-3-phosphate dehydrogenase (EC
1.2.1.12).
GN GAPA OR SMU.360.

OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian Y.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.,
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
DR EMBL; AB014883; AAN58118.1; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHHRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 337 AA; 36068 MW; 42BFE20365963C22 CRC64;

Query Match 92.1%; Score 1579.5; DB 16; Length 337;
Best Local Similarity 90.2%; Pred. No. 6.9e-98;
Matches 304; Conservative 18; Mismatches 14; Indels 1; Gaps 1;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFEVNGFKVSAERDPENIDWATDGVVEIVLEATGFFAKKEAEKHLHAN-GAKKV 119
Db 61 KEGGFEVNGFKVSAERDPENIDWATDGVVEIVLEATGFFAKKEAEKHLHAN-GAKKV 120

QY 120 ITAPGGNDVKTIVFNTNHDILDTGTETVISGASCTNCLAPMAKALHDAFGIQGLMTTIH 179
Db 121 ITAPGGNDIKTIVFNTNHDVLDGTETVISGASCTNCLAPMAKALHDFSIKEGLMTTIH 180

QY 180 AYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVP 239
Db 181 AYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVP 240

QY 240 TGSVTELVVTLTKNVSVDENAMKAASNDSTFGYTEDPIVSSDIVGVSGLFDATQTKV 299
Db 241 TGSVTELVAVLDKKVTVDENAMKAANESYGYTEDPIVSSDIVGMSFGSLFDATQTKV 300

QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 LDVDGKQLVKVSWYDNEMSYTSQLVRLTLEYFAKIAK 337

RESULT 11
Q8KVU5
ID Q8KVU5 PRELIMINARY; PRT; 336 AA.
AC Q8KVU5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus parauberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1348;
RN [1]
RP SEQUENCE FROM N.A.
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;

RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421901; AAM73772.1; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHHRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36037 MW; 311647C25489AC9E CRC64;

Query Match 92.0%; Score 1578; DB 2; Length 336;
Best Local Similarity 90.8%; Pred. No. 8.6e-98;
Matches 305; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFEVNGFKVSAERDPENIDWATDGVVEIVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGFKVSAERDPENIDWATDGVVEIVLEATGFFAKKEAEKHLHANGAKKVI 120

QY 121 TAPGGNDVKTIVFNTNHDILDTGTETVISGASCTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 121 TAPGGDDVKTIVFNTNHDILDTGTETVISGASCTNCLAPMAKALQDNFVGKQGLMTTIHA 180

QY 181 YTGDMILDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVP 240
Db 181 YTGDMILDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVP 240

QY 241 GSVTELVVTLTKNVSVDENAMKAASNDSTFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVAVLNKETSVEEINSMKAAANDSYGYTEDPIVSSDIVGMSFGSLFDATQTKV 300

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTAQLDRTLEYFAKIAK 336

RESULT 12
Q8VVB9
ID Q8VVB9 PRELIMINARY; PRT; 336 AA.
AC Q8VVB9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
GN GAPDH.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMG18311;
RA van den Bogaard P.T.C., Kleerebezem M., Hols P., Crispie F.,
RA Kuipers O.P., de Vos W.M.;
RT "Modulation of Glycolysis by Lactose Availability in Streptococcus
RT thermophilus";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF442551; AAL35377.1; -;

```
DR HSSP; P00354; 3GPD.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;

Query Match
Best Local Similarity 91.7%; Score 1572; DB 2; Length 336;
Matches 303; Conservative 18; Mismatches 15; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 120
Db 61 KEGGFVNGNFIKVSARDPENIDWATDGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 120

QY 121 TAPGGNDVKTIVFNTNHDILDTGTEVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180
Db 121 TAPGGNDVKTIVFNTNHDILDTGTEVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA 180

QY 181 YTGDMILDPGHRGGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPGHRGGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

QY 241 GSVTELVVTLTKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKVM 300

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

QY 301 DVDGKQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 DVDGKQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 13
Q9L5X6
ID Q9L5X6 PRELIMINARY; PRT; 320 AA.
AC Q9L5X6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Extracellular glyceraldehyde-3-phosphate dehydrogenase
DE (Fragment).
OS Streptococcus gordonii.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1302;
RN [1]
RP SEQUENCE FROM N.A.
RA Nelson D., Goldstein J.M., Boatright K., Harty D.W.S., Cook S.L.,
RA Hickman P.J., Potempa J., Travis J., Mayo J.A.;
RT "Purification and characterization of an extracellular
RT glyceraldehyde-3-phosphate dehydrogenase from Streptococcus sanguis
RT and cloning of the gene encoding this enzyme."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF247678; AAF64063.1; -.
DR HSSP; P00362; 1GD1.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
```

```
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1 320
FT NON_TER 320
SQ SEQUENCE 320 AA; 33968 MW; 6FFCC18BB01E91F0 CRC64;

Query Match
Best Local Similarity 88.5%; Score 1518; DB 2; Length 320;
Matches 292; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 4 KVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEVKEG 63
Db 1 KVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEVKEG 60

QY 64 GFEVNGNFIKVSARDPENIDWATDGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 123
Db 61 GFEVNGNFIKVSARDPENIDWATDGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 120

QY 124 GENDVKTIVFNTNHDILDTGTEVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHAYTG 183
Db 121 GSDVKTIVFNTNHDVLDGTEVISGASCTTNCLAPMAKALQDNFVGLMTTIHAYTG 180

QY 184 QMILDPGHRGGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGSV 243
Db 181 QMILDPGHRGGDLRRARAGAAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGSV 240

QY 244 TELVTLTKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKVMEVD 303
Db 241 TELVTLTKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDTATQTKVMEVD 300

QY 304 GSQLVKVSWSYDNEMSYTAQ 323
Db 301 GKQLVKVSWYDNEMSYTAQ 320

RESULT 14
Q9AJT7
ID Q9AJT7 PRELIMINARY; PRT; 309 AA.
AC Q9AJT7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
DE GAP.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=clinical isolate;
RA Amezaga M.R., Carter P.E., Cash P., McKenzie H.;
RT "Sequence variation in mef(A) and gap genes in M-phenotype
RT erythromycin resistant Streptococcus pneumoniae."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AJ292048; CAC27448.1; -.
DR HSSP; P00362; 1GD1.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
```

DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 309
SQ SEQUENCE 309 AA; 32930 MW; 89E45C3BEA6A0528 CRC64;

Query Match 85.6%; Score 1468.5; DB 2; Length 309;
Best Local Similarity 91.9%; Pred. No. 1.7e-90;
Matches 285; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 9 GFGRIGRLAFRRIONVEGVETRIINDLTDPNMLAHLKYYDTTQGRFDGTVVEVKEGGFEVN 68
Db 1 GFGRIGRLAFRRIONVEGVETRIINDLTDPVMLAHLKYYDTTQGRFDGTVVEVKEGGFEVN 60

QY 69 GNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVVITAPGGNDV 128
Db 61 GKFIKVSARDPEQIDWATDGVIEIVLEATGFFAKKEAAEKHL-KGGAKKVVITAPGGNDV 119

QY 129 KTVVFNTHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAYTGDQMIL 188
Db 120 KTVVFNTHDVLDTGTETVISGASCTTNCLAPMAKALQDNFVGVVEGLMTTIHAYTGDQMIL 179

QY 189 DGPHRGDLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGVSVELV 248
Db 180 DGPHRGDLRRARAGAANIVPNSTGAAKAIGLVIPELNGKLDGSAQRVPTPTGVSVELVA 239

QY 249 TLDKNVSVDEINAAKKAASNDSEFGYTEDPIVSSDIVGVSYSGLFDTQTCKWMEVDGSQLV 308
Db 240 VLEKNVTVDENVNAAMKAASNESYGYTEDPIVSSDIVGMSYGSFLFDTQTCKVLDVDGKQLV 299

QY 309 KVSWSYDNEM 318
Db 300 KVSWSYDNEM 309

Search completed: May 11, 2004, 16:49:18
Job time : 52 secs

RESULT 15

Q9AJT9 PRELIMINARY; PRT; 309 AA.

ID Q9AJT9
AC Q9AJT9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
GN GAP.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=clinical isolate;
RA Amezaga M.R., Carter P.E., Cash P., McKenzie H.;
RT "Sequence variation in mef(A) and gap genes in M-phenotype
erythromycin resistant Streptococcus pneumoniae."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
DEHYDROGENASE FAMILY.
DR EMBL; AJ292046; CAC27446.1; -.
DR HSSE; P00362; 1GD1.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 309

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:49:42 ; Search time 48 Seconds
(without alignments)
1942.967 Million cell updates/sec

Title: US-10-650-369-12
Perfect score: 1715
Sequence: 1 MVVKVINGFRIGRLAFRR.....EMSYTAQLVRLTYFAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues 1140673

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1715	100.0	336	9	US-09-878-766A-12
2	1715	100.0	336	10	US-09-878-781-4
3	1715	100.0	336	12	US-10-650-369-12
4	1715	100.0	336	14	US-10-134-297-4
5	1714	99.9	336	12	US-10-282-122A-74379
6	1688.5	98.5	335	10	US-09-878-781-14
7	1656.5	96.6	448	9	US-09-878-766A-22
8	1656.5	96.6	448	12	US-10-650-369-22
9	1618	94.3	336	9	US-09-878-766A-16
10	1618	94.3	336	10	US-09-878-781-8
11	1618	94.3	336	12	US-10-650-369-16
12	1618	94.3	336	14	US-10-134-297-8
13	1611	93.9	336	10	US-09-878-781-16
14	1602	93.4	336	9	US-09-878-766A-14
15	1602	93.4	336	10	US-09-878-781-6

16	1602	93.4	336	12	US-10-650-369-14	Sequence 14, Appl
17	1602	93.4	336	14	US-10-134-297-6	Sequence 6, Appli
18	1593.5	92.9	359	9	US-09-815-242-13169	Sequence 13169, A
19	1593.5	92.9	359	12	US-10-282-122A-74186	Sequence 74186, A
20	1592.5	92.9	359	9	US-09-815-242-13593	Sequence 13593, A
21	1584	92.4	336	9	US-09-878-766A-20	Sequence 20, Appl
22	1584	92.4	336	10	US-09-878-781-12	Sequence 12, Appl
23	1584	92.4	336	12	US-10-650-369-20	Sequence 20, Appl
24	1584	92.4	336	14	US-10-134-297-12	Sequence 12, Appl
25	1579.5	92.1	337	12	US-10-282-122A-72254	Sequence 72254, A
26	1578	92.0	336	9	US-09-878-766A-18	Sequence 18, Appl
27	1578	92.0	336	10	US-09-878-781-10	Sequence 10, Appl
28	1578	92.0	336	12	US-10-650-369-18	Sequence 18, Appl
29	1578	92.0	336	14	US-10-134-297-10	Sequence 10, Appl
30	1549.5	90.3	333	9	US-09-765-272-54	Sequence 54, Appl
31	1410	82.2	336	15	US-10-369-493-18524	Sequence 18524, A
32	1338.5	78.0	337	15	US-10-369-493-18342	Sequence 18342, A
33	1337.5	78.0	333	9	US-09-815-242-10847	Sequence 10847, A
34	1337.5	78.0	333	12	US-10-282-122A-57294	Sequence 57294, A
35	1277	74.5	335	12	US-10-282-122A-53266	Sequence 53266, A
36	1262	73.6	336	12	US-10-282-122A-60492	Sequence 60492, A
37	1259.5	73.4	334	12	US-10-282-122A-51519	Sequence 51519, A
38	1217.5	71.0	335	12	US-10-282-122A-71734	Sequence 71734, A
39	1197.5	69.8	334	12	US-10-282-122A-65661	Sequence 65661, A
40	1196.5	69.8	357	12	US-10-282-122A-65001	Sequence 65001, A
41	1190	69.4	336	12	US-10-282-122A-70890	Sequence 70890, A
42	1183	69.0	336	9	US-09-815-242-5618	Sequence 5618, Ap
43	1183	69.0	336	9	US-09-815-242-12540	Sequence 12540, A
44	1183	69.0	336	12	US-10-282-122A-43997	Sequence 43997, A
45	993	57.9	333	12	US-10-282-122A-56828	Sequence 56828, A

ALIGNMENTS

RESULT 1
US-09-878-766A-12
; Sequence 12, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-766A-12

Query Match	100.0%;	Score 1715;	DB 9;	Length 336;
Best Local Similarity	100.0%;	Pred. No. 3.1e-154;		
Matches 336;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MVVKVINGFRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV	60	
Db	1	MVVKVINGFRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV	60	
QY	61	KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAEKHLHANGAKKVI	120	
Db	61	KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAEKHLHANGAKKVI	120	
QY	121	TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA	180	
Db	121	TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQGLMTTIHA	180	
QY	181	YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT	240	

Db 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDEINAAKAAASNDSPFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDEINAAKAAASNDSPFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 2
US-09-878-781-4
; Sequence 4, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-781-4

Query Match 100.0%; Score 1715; DB 10; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTIVFNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGGNDVKTIVFNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIIA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDEINAAKAAASNDSPFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDEINAAKAAASNDSPFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 3
US-10-650-369-12
; Sequence 12, Application US/10650369
; Publication No. US20040062774A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650,369
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878,766A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-10-650-369-12

Query Match 100.0%; Score 1715; DB 12; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTIVFNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGGNDVKTIVFNTNHDILDTGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIIA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDEINAAKAAASNDSPFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDEINAAKAAASNDSPFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 4
US-10-134-297-4
; Sequence 4, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-10-134-297-4

Query Match 100.0%; Score 1715; DB 14; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120

Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTVVNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGNDVKTVVNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
QY 181 YTGDMILDPHRRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPHRRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDENIAAMKAASNDSTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 300
Db 241 GSVTELVVTLTKNVSVDENIAAMKAASNDSTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 5
US-10-282-122A-74379
; Sequence 74379, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 74379
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74379

Query Match 99.9%; Score 1714; DB 12; Length 336;
Best Local Similarity 99.7%; Pred. No. 3.8e-154;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTVVNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGNDVKTVVNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
QY 181 YTGDMILDPHRRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPHRRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDENIAAMKAASNDSTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 300
Db 241 GSVTELVVTLTKNVSVDENIAAMKAASNDSTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 6
US-09-878-781-14
; Sequence 14, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 14
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SpyGapC
; OTHER INFORMATION: protein
US-09-878-781-14

Query Match 98.5%; Score 1688.5; DB 10; Length 335;
Best Local Similarity 98.8%; Pred. No. 1e-151;
Matches 332; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 119
QY 121 TAPGGNDVKTVVNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGNDVKTVVNTNHDILDTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 179
QY 181 YTGDMILDPHRRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPHRRGGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 239
QY 241 GSVTELVVTLTKNVSVDENIAAMKAASNDSTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 300
Db 241 GSVTELVVTLTKNVSVDENIAAMKAASNDSTGTTVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 300

Db 240 GSVTELVVTLTKNSVDEINSAMKAASNDSEFGYTEDPIVSSDIVGSYSLFDATQTKVM 299
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 300 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 7

US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

Query Match 96.6%; Score 1656.5; DB 9; Length 448;
Best Local Similarity 79.6%; Pred. No. 1.7e-148;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
QY 1 MVVKVGINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 60
Db 28 MVVKVGINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 87
QY 61 KEGGFVNGNFIKVSARDPE----- 81
Db 88 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEALEGTVKGGFVNGKFIKVSAAE 147
QY 82 -----NIDWATDGVETVLE 95
Db 148 KDEQIDWATDGVETVLEIDGTVEVKEGGFVNGQFVKVSAEREPAIDWATDGVETVLE 207
QY 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVVISGASCTTN 155
Db 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVVISGASCTTN 267
QY 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 215
Db 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 327
QY 216 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVVTLTKNSVDEINAAKKAASNDSEFGYTE 275
Db 328 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVVTLTKNSVDEINAAKKAASNDSEFGYTE 387
QY 276 DPIVSSDIVGSYSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
Db 388 DPIVSSDIVGSYSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
QY 336 K 336
Db 448 K 448

RESULT 8

US-10-650-369-22
; Sequence 22, Application US/10650369
; Publication No. US20040062774A1
; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650,369
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878,766A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-10-650-369-22

Query Match 96.6%; Score 1656.5; DB 12; Length 448;
Best Local Similarity 79.6%; Pred. No. 1.7e-148;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
QY 1 MVVKVGINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 60
Db 28 MVVKVGINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKDYDTTQGRFDGTVEV 87
QY 61 KEGGFVNGNFIKVSARDPE----- 81
Db 88 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEALEGTVKGGFVNGKFIKVSAAE 147
QY 82 -----NIDWATDGVETVLE 95
Db 148 KDEQIDWATDGVETVLEIDGTVEVKEGGFVNGQFVKVSAEREPAIDWATDGVETVLE 207
QY 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVVISGASCTTN 155
Db 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVVISGASCTTN 267
QY 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 215
Db 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 327
QY 216 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVVTLTKNSVDEINAAKKAASNDSEFGYTE 275
Db 328 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVVTLTKNSVDEINAAKKAASNDSEFGYTE 387
QY 276 DPIVSSDIVGSYSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
Db 388 DPIVSSDIVGSYSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
QY 336 K 336
Db 448 K 448

RESULT 9

US-09-878-766A-16
; Sequence 16, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match      94.3%;   Score 1618;   DB 9;   Length 336;
Best Local Similarity 92.9%;   Pred. No. 5.1e-145;
Matches 312;   Conservative 15;   Mismatches 9;   Indels 0;   Gaps 0;

QY      1  MVVKVGINFGFRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db      1  MVVKVGINFGFRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60

QY      61  KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db      61  KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY      121  TAPGGNDVKTIVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db      121  TAPGGDDVKTIVFNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180

QY      181  YTGDMILDGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db      181  YTGDMILDGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

QY      241  GSVTELVVTLDKNVSVDENINAAKASNDSPGYTEDPIVSSDIVGVSGLFDTATQTKVM 300
Db      241  GSVTELVVTLDKNVSVDENINAAKASNDSPGYTEDPIVSSDIVGVSGLFDTATQTKVM 300

QY      301  EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db      301  TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
```

```
RESULT 10
US-09-878-781-8
; Sequence 8, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-781-8

Query Match      94.3%;   Score 1618;   DB 10;   Length 336;
Best Local Similarity 92.9%;   Pred. No. 5.1e-145;
Matches 312;   Conservative 15;   Mismatches 9;   Indels 0;   Gaps 0;

QY      1  MVVKVGINFGFRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db      1  MVVKVGINFGFRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60

QY      61  KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db      61  KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY      121  TAPGGNDVKTIVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db      121  TAPGGDDVKTIVFNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180
```

```
QY      181  YTGDMILDGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db      181  YTGDMILDGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

QY      241  GSVTELVVTLDKNVSVDENINAAKASNDSPGYTEDPIVSSDIVGVSGLFDTATQTKVM 300
Db      241  GSVTELVVTLDKNVSVDENINAAKASNDSPGYTEDPIVSSDIVGVSGLFDTATQTKVM 300

QY      301  EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db      301  TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
```

```
RESULT 11
US-10-650-369-16
; Sequence 16, Application US/10650369
; Publication No. US20040062774A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650,369
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878,766A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-10-650-369-16
```

```
Query Match      94.3%;   Score 1618;   DB 12;   Length 336;
Best Local Similarity 92.9%;   Pred. No. 5.1e-145;
Matches 312;   Conservative 15;   Mismatches 9;   Indels 0;   Gaps 0;

QY      1  MVVKVGINFGFRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db      1  MVVKVGINFGFRIGRLAFRRIONVEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60

QY      61  KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db      61  KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY      121  TAPGGNDVKTIVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db      121  TAPGGDDVKTIVFNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFVGKQGLMTTIHA 180

QY      181  YTGDMILDGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db      181  YTGDMILDGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

QY      241  GSVTELVVTLDKNVSVDENINAAKASNDSPGYTEDPIVSSDIVGVSGLFDTATQTKVM 300
Db      241  GSVTELVVTLDKNVSVDENINAAKASNDSPGYTEDPIVSSDIVGVSGLFDTATQTKVM 300

QY      301  EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db      301  TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
```

```
RESULT 12
US-10-134-297-8
; Sequence 8, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
```

```
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-10-134-297-8

Query Match      94.3%; Score 1618; DB 14; Length 336;
Best Local Similarity 92.9%; Pred. No. 5.1e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY      1  MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
Db      1  MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60

QY      61  KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db      61  KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY      121  TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db      121  TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180

QY      181  YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db      181  YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

QY      121  TAPGGDDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFGVKQGLMTTIHA 180
Db      121  TAPGGDDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFGVKQGLMTTIHA 180

QY      181  YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db      181  YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

QY      241  GSVTELVVTLTKNVSVDENIAAMKAASNDSPGYTEDPIVSSDIVGSYGSFLDATQTQKV 300
Db      241  GSVTELVVTLTKNVSVDENIAAMKAASNDSPGYTEDPIVSSDIIGMAYGSFLDATQTQKV 300

QY      301  EVDGSQLVKVSWYDNEMSYTAQLVRLTYEYFAKIAK 336
Db      301  TVDGNQLVKVSWYDNEMSYTAQLVRLTYEYFAKIAK 336
```

RESULT 13

```
US-09-878-781-16
; Sequence 16, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SeqGapC
; OTHER INFORMATION: protein
US-09-878-781-16

Query Match      93.9%; Score 1611; DB 10; Length 336;
Best Local Similarity 95.2%; Pred. No. 2.3e-144;
Matches 320; Conservative 2; Mismatches 14; Indels 0; Gaps 0;
```

```
QY      1  MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
```

```
Db      1  MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60

QY      61  KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db      61  KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY      121  TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db      121  TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180

QY      181  YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db      181  YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

QY      241  GSVTELVVTLTKNVSVDENIAAMKAASNDSPGYTEDPIVSSDIVGSYGSFLDATQTQKV 300
Db      241  GSVTELVVTLTKNVSVDENIAAMKAASNDSPGYTEDPIVSSDIVGSYGSFLDATQTQKV 300

QY      301  EVDGSQLVKVSWYDNEMSYTAQLVRLTYEYFAKIAK 336
Db      301  EVDGSQLVKVSWYDNEMSYTAQLVRLTYEYFAKIAK 336

RESULT 14
US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match      93.4%; Score 1602; DB 9; Length 336;
Best Local Similarity 91.7%; Pred. No. 1.7e-143;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;
```

```
QY      1  MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
Db      1  MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60

QY      61  KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db      61  KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY      121  TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db      121  TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFGVKQGLMTTIHA 180

QY      181  YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db      181  YTGDMILDGPHRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

QY      241  GSVTELVVTLTKNVSVDENIAAMKAASNDSPGYTEDPIVSSDIVGSYGSFLDATQTQKV 300
Db      241  GSVTELVATLEKDVVEEVNAAKAAANDSYGYTEDPIVSSDIVGSYGSFLDATQTQKV 300

QY      301  EVDGSQLVKVSWYDNEMSYTAQLVRLTYEYFAKIAK 336
Db      301  TVDGNQLVKVSWYDNEMSYTAQLVRLTYEYFAKIAK 336
```


RESULT 15
US-09-878-781-6
; Sequence 6, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-781-6

Query Match 93.4%; Score 1602; DB 10; Length 336;
Best Local Similarity 91.7%; Pred. No. 1.7e-143;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIQLAFRRIONVEGVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIQLAFRRIONVEGVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTVEV 60

QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGQFVKVSAEREPANIDWATDGVIEIVLEATGFFASKEKAGQHIHENGAKKVI 120

QY 121 TAPGGNDVKTVEFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGNDVKTVEFNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFQVKGGLMTTIHA 180

QY 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240

QY 241 GSVTELVVTLTKNVSVDEINAAKAAANDSFGYTEDPIVSSDIVGSYGSLEFDATQTKVM 300
Db 241 GSVTELVATLEKDVTVVEEVNAAKAAANDSYGYTEDPIVSSDIVGSYGSLEFDATQTKVQ 300

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSQLVRLLEYFAKIAK 336

Search completed: May 11, 2004, 16:55:20
Job time : 49 secs